On Corruption and Illegal Strategies

Esther Brügger

Inauguraldissertation zur Erlangung der Würde eines
DOCTOR RERUM OECONOMICARUM
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Universität Bern

2004

Die Fakultät hat diese Arbeit am 17. Juni 2004 auf Antrag der beiden Gutachter Prof. Dr. Gunter Stephan und Prof. Dr. Aleksander Berentsen als Dissertation angenommen, ohne damit zu den darin ausgesprochenen Auffassungen Stellung nehmen zu wollen.

Preface

Corruption is an age-old phenomenon and a widespread, multifaceted problem: there is hardly any country that has been spared from scandals in the past and corruption has appeared in almost any kind of interaction between the private and the public sector.

In 1995, the Sunday Times of India published a "bribe index", i.e. a list of the customary bribes that had to be paid for a range of routine public services. For example, for the issuing of a driver's license, a bribe of between 1,000 and 2,000 rupies was asked; For the installation of an electric meter, government officials usually pocketed between 25,000 and 30,000 rupies.

After a large department store collapsed in Korea in 1995, it was discovered that the accident had been caused by the substandard concrete that had been used in construction. Government officials had taken bribes to allow violation of the safety standards, which reduced the building costs significantly for the department store.

In Gambia, people bribe tax collectors in order to reduce their duty. In the early nineties, the country's forgone tax revenue amounted to 8 - 9 % of GDP. This was about eight times the country's spending on health.

When Italy conducted an anticorruption investigation in 1991, the construction costs for the Milan subway fell from $227 million to $97 million per kilometer. Corrupt officials had agreed to overpriced contracts with construction companies, taking bribes as return services.[1]

Empirical economic research has been concerned with the *consequences* as well as the *causes* of corruption. The most important consequences are the following: Corruption increases poverty and income inequality,[2] it augments

---

[1] All examples from Rose-Ackerman (1999).

[2] Gupta et al. (2002)

the public deficit and public debt,[3] it reduces growth and investment,[4] the effectiveness of social spending, the formation of human capital,[5] expenditure on education and health,[6] and tax revenue[7]. Cross-country studies concerned with the causes of a society's corruption level identify income as the single most important determinant. Cultural and political factors are not crucial for a country's corruption rate.[8]

There is a long tradition of microeconomic models that deal with the individual incentives of engaging in corrupt activities. However, the consequences of corruption are measured at the societal level. In fact, very little research has been conducted to provide the missing link between the causes at the individual level and the consequences at the societal level.[9] To comprehend both levels within the same model is important because it is beyond controversy that the consequences of corruption have an influence on the individual's incentives. Therefore, only a dynamic setup is suited to capturing all relevant factors of corruption. This is the motivation for the first chapter of this book, which extends standard evolutionary game theory to a new class of games allowing us to study the feedback effects of population variables on individual incentives. In particular, the model includes the informational distortions prevailing in the case of corruption.

While the first chapter applies a shortcut for modelling the information available to the individuals involved, the second chapter addresses the issue of information. It explicitly describes the relation between the nature of information and the spread of illegal activities in a population. We abandon the specific formulation of corruption used in the previous chapter and comprehend it as a general illegal activity. By extending a standard spatial evolutionary game with heterogeneity of agents and various information settings, we can explore the imitation dynamics peculiar to the spread of illegal activities.

The third chapter extends the analysis of the second. Numerical simulations of the formally described model with local information and heterogeneous agents allow us to draw conclusions about contagion and spread of illegal activities under general conditions.

[3] Tanzi (2002)

[4] Mauro (1995); Knack and Keefer (1995)

[5] Gupta et al. (2002)

[6] Tanzi (2002)

[7] Tanzi and Davoodi (2000)

[8] Paldam (2002)

[9] Exceptions are Chakrabarti (2001) and Andvig and Moene (1990).

Acknowledgement

I would like to express my gratitude to my advisor, Gunter Stephan, for his support and his encouragement throughout my graduate studies.

I am deeply indebted to Aleksander Berentsen and Simon Lörtscher. The second chapter of this book has benefited greatly from our joint work on Research Project B-61, WWZ-Forum University of Basel. Their editorial advice was essential to the completion of the second chapter. Our collaboration has taught me innumerable lessons and insights on the workings of academic research in general.

I have greatly benefited from the Brown-Bag seminars at the Economics Department in Bern and the Max Planck Institute for Research into Economic Systems in Jena.

I thank Thomas Brenner, Alain Egli, Werner Güth, Armin Hartmann, Roland Hodler, Michael Manz, and Manuel Wälti for valuable comments and suggestions. I am grateful to Jean-Jacques Laffont for motivating me to study corruption.

Many thanks go to Kurt Schmidheiny for advice on programming the Matlab codes, and to Andres Aeschlimann at the Informatikdienste, University of Bern, for providing me access to the large-capacity computer for my simulations. I am grateful to Peter Mani-Levitska for having taken time to discuss my questions concerning the nonlinear dynamical systems.

I gratefully acknowledge financial support by the WWZ Forum Basel.

My parents, Helene and Ernst, receive my deepest gratitude for their dedication and the many years of support during my undergraduate studies that provided the foundation for this work.

Contents

Preface **v**

Contents **ix**

1 The Dynamics of Corruption **1**
1.1 Introduction 1
1.2 The Model 8
1.2.1 The Standard Framework 8
1.2.2 The Extension of the Standard Framework 17
1.3 Frequency Dependent Evolutionary Games 18
1.3.1 Some General Results 19
1.3.2 The Corruption Game 21
1.4 Conclusions 33
1.A Appendix: Proofs 36
1.A.1 Proof of Proposition 1 36
1.A.2 Proof of Proposition 2 37
1.A.3 Proof of Proposition 3 37
1.A.4 Proof of Proposition 4 39
1.A.5 Proof of Proposition 5 40
1.B Appendix: Games with two Strategies 56

2 Imitating Illegal Activities **61**
2.1 Introduction 61
2.2 Preliminary Considerations 62
2.2.1 Imitation as Social Interaction 62
2.2.2 Informational Distortions 63
2.2.3 The Theoretical Approach 64

2.3 Homogeneity of Agents . . . 68
2.3.1 The Stage Game . . . 68
2.3.2 Location and Imitation . . . 69
2.3.3 Absorbing States . . . 71
2.3.4 The Role of Information . . . 72
2.4 Heterogenous Agents with Global Information . . . 73
2.4.1 Asymmetric Games . . . 74
2.4.2 The Role of Information . . . 75
2.5 Heterogenous Agents with Local Information . . . 76
2.5.1 The Role of Information . . . 77
2.5.2 Maximal Segregation . . . 80
2.5.3 Minimal Segregation of Types . . . 83
2.5.4 Maximal versus Minimal Segregation . . . 85
2.6 Extensions . . . 88
2.6.1 Mutations . . . 88
2.6.2 Gradual Heterogeneity and Effectiveness of D . . . 91
2.7 Conclusions . . . 93
2.A Appendix: Proofs . . . 96
2.A.1 Proof of Proposition 3 . . . 96
2.A.2 Proof of Proposition 4 . . . 97
2.A.3 Decision Terms for Edge Players . . . 99
2.A.4 Decision Terms for Double-Edge Players . . . 100
2.A.5 Proof of Proposition 5 . . . 100
2.A.6 Proof of Proposition 6 . . . 105

3 Informational Deficits and Strategy Clustering 107
3.1 Introduction . . . 107
3.2 Monte Carlo Simulations . . . 109
3.2.1 The Procedure . . . 109
3.2.2 Standard Simulation Settings . . . 111
3.3 Characterizing the Absorbing States . . . 111
3.3.1 Discerning Strategy Choices of H- and L-Types . . . 112
3.3.2 Different Positions in the Type Distribution . . . 113
3.3.3 Strategy Clusters . . . 116
3.4 Varying the Standard Settings . . . 119
3.4.1 Varying the Initial Strategy State . . . 119

3.4.2 Varying the Type Distribution . . . 122
3.5 Conclusions . . . 123
3.A Appendix: An Example of the Evolutionary Game . . . 125

Notation 127

Bibliography 133

Chapter 1

The Dynamics of Corruption

> The efforts of men are utilized in two different ways: they are directed to the production or transformation of economic goods, or else to the appropriation of goods produced by others.
>
> Vilfredo Pareto

1.1 Introduction

Corruption is the major hindrance for economic growth and overall development in poverty-stricken countries. The harmful impact of corruption has several reasons. In a corrupt environment private investment, both domestic and foreign, yields lower returns because of additional costs and a climate of heightened uncertainty. The resulting decreased investment incentives lower growth (see Mauro, 1995, for empirical evidence). Not only is the investment level sub-optimal in countries with prevalent corruption, but also the management of public services. Corruption raises the prices of public services,[1] drives up the concerning transaction costs,[2] and leads to general mismanagement in the public sector (Rose-Ackerman, 1978; Shleifer and Vishny, 1993). The reason for the latter is that corrupt policy-makers

[1] One reason for high prices is that policy makers are bribed to agree on overpriced offers and contracts with private companies (Rose-Ackerman, 1999).

[2] Administration and government become less efficient, because bribe recipients shorten the supply of services or invent new corruption opportunities by extending and complicating the administrative process (Rose-Ackerman, 1999).

allocate public funds inefficiently, seeking to direct money into projects that easily can be corrupted. Therefore, a corrupt government and a corrupt administration leave the society's optimal intertemporal spending path (Jain, 2001).[3] Low growth rates and inefficient allocation of public funds lead to persistent poverty. Furthermore, corruption raises income inequality through other channels such as an inaccessible education system and a tax regime in favor of those who can afford shirking their duties. Additionally, a society's most disadvantaged groups are negatively affected by the fact that foreign aid projects seldom bring about the desired aim: In a corrupt environment resources seep away before they reach the needy.

Corruption rates[4] vary considerably across countries and time. It is our aim to contribute to the understanding of these observed differences. In particular we are interested in determining the conditions that lead to a corruption rate which prevents economic growth and development of a society. We believe that a thorough understanding of the dynamics of corruption is an inevitable prerequisite to form policy recommendations that can improve present situations.

The corruption rate is the outcome of individual choices to engage in corrupt activities (Chakrabarti, 2001). Consequently, we address the problem on the individual level and model the incentives that lead individuals to corrupt activities explicitly. Analyzing corruption on the microeconomic level has been the most traditional approach in studying the phenomenon in economic theory.[5] Our analysis differs from previous theoretical research on corruption in two respects. Firstly, we explicitly model the informational complicacies individuals face when making the decision to adopt or not to adopt corrupt behavior. Our approach to this special informational situation in the case of corruption is evolutionary game theory. Within this framework we specify an underlying game that represents the interaction of corrupt and fair behavior in a society. Secondly, we include the effects that

[3]One example for the distortion in public expenditure through corruption is treated in Mauro (1997): Corruption leads to a negative relation between the corruption level in a society and its spending for education and health.

[4]It surpasses the purpose of this paper to specify or discuss the ways and means of measuring corruption. The term *corruption rate* is used as a measure for the degree to which corruption exists among public officials and politicians in a country.

[5]Theoretical research dates back to the 1970s with Krueger (1974) and Rose-Ackerman (1978) and is now included in many different fields in economics.

resulting population behavior has on the decision of the single individual.[6] If two factors in a model have mutual impact on each other in a dynamic context, we say that a *feedback effect* exists between them. We will not only include the feedback effects between the corruption rate and the individual decision to adopt corrupt behavior in our consideration, but also the feedback effect between the individual decision and per capita income.

The consideration of feedback effects requires an extension of standard evolutionary game theory. We therefore define the class of *frequency dependent evolutionary games* and deliver some general results necessary for their analysis. In the following we substantiate the choice of evolutionary game theory as our framework and explain why we think that it suits the special informational situations where corruption decisions are made. Based on this we address the issue of the most important feedback effects surrounding corruption and describe how we extend standard evolutionary game theory to seize these feedback effects.

In our view it seems natural to study the spread of illegal activities such as corruption in the framework of evolutionary game theory. If an activity is illegal and can be prosecuted, respective information is scarce. The reason is that individuals are not willing to share related information because this reveals their relation to the illegal activity. Therefore, individuals actually may be completely oblivious to the existence of a strategy, may not know what share of the population plays the strategy, what the strategy's returns are, or simply how to play the strategy.

Individuals acting in situations where information is too scarce for beliefs to be formed and circumstances to be fully comprehended, are ideally represented by models of bounded rationality (Simon, 1955, 1987; Selten, 2001; Dequech, 2001). The reason is that we expect individuals to act myopically (the behavior of others cannot be anticipated) and with inertia (gathering information is difficult and time consuming as a result). Myopic behavior in combination with inertia defines bounded rationality (Kandori et al., 1993). Imitation is one way of learning how to play a game in a situation described above (see Cartwright, 2002, for a survey of imitation behavior rules). An evolutionary game combines myopic behavior, inertia

[6]Andvig and Moene (1990) models the impact of the corruption rate on individual choice. What we have in mind, however, is a more comprehensive model, including feedback effects between per capita income and corruption.

and imitation in a suitable way for our purposes.[7]

The structure of the evolutionary game is that a continuum of infinitely-lived players chooses between corrupt and fair strategies of an underlying game. They are then matched pairwise to play the underlying game. Players make their strategy decision according to an imitation rule which includes inertia and myopic behavior.

Thereby the imitation rule determines the decisions of the players: in every period of the game each player observes the strategy and the expected payoff of another player drawn at random. Consequently, a frequently played strategy will be observed with a greater probability than a strategy played by a minority. Players then imitate the observed strategy if the expected payoff of their own strategy is lower than the one of the observed.

The strategy decisions of the whole population define the adjustment dynamics which in our case coincide with the replicator dynamics. The adjustment dynamics is formalized as a system of ordinary differential equations and can be analyzed by determination of the evolutionary equilibria and description of the solutions' global behavior. Let us now turn to the underlying game which describes the incentives for corrupt behavior and the feedback effects named above.

Empirical research has studied the correlations between the corruption rate in a society and both cultural and economic variables (see Jain, 2001; Graf Lambsdorff, 1999, for surveys of the empirical research). From the many interesting and significant correlations that have been isolated, per capita income has proven to be the most reliable predictor of the corruption rate over diverse data sets: There is a strong negative correlation between the corruption rate and GDP per capita of a country in most empirical studies.[8]

The direction of causality between income and corruption rate has been

[7]To be specific: We apply a continuous-time pure-strategy selection dynamics arising from strategy adaptation by myopic imitation.

[8]Mauro (1995) is using data consisting of the Business International (BI) indices on corruption, red tape, and efficiency of the judicial system; Husted (1999) and Treisman (2000) take the Transparency International's (TI) annual index of perceived corruption as a dependent variable; Paldam (2002) as well relies on the TI CPI index. For a nontechnical discussion see Bardhan (1997).

more difficult to tackle than the correlation between the two. There are sound theoretical arguments for both directions of impact. Corruption affects GDP per capita by lowering private investment and thereby lowering growth (Mauro, 1995).[9] Lowered institutional quality because of corruption is another reason for decreased growth rates.[10] While this direction of causality is well established in the literature, we also find research supporting the reverse causality. Property rights are more valuable in high income countries. Economic reasoning suggests that people are then willing to spend more to protect property rights (Eggertsson, 1990) which increases the quality of institutions. We also expect the opportunity costs of corrupt behavior to be higher because return to investment and labor is high. While we often lack a time series long enough to highlight the hypothesis that an economy has significantly reduced its corruption rate when having undergone economic development, there are many historical examples describing this process (see Hofstaedter, 1973, for examples). In order to shed light on the matter of causality between economic growth and the corruption rate, Chong and Calderon (2000) choose the following approach. They apply a method (Geweke, 1979) based on Granger causality to corruption data, which allows to measure linear feedbacks. They find that causality works in both directions: institutional quality causes growth *and* growth causes institutional quality.

How can we include the result of Chong and Calderon (2000) in our model? In order to model the impact of corruption on income, we define the underlying game in the following way. The underlying game consists of three strategies: agents choose between a private sector activity, being a fair government employee, and being a corrupt government employee. The private activity's return is a fixed surplus when playing against a fair government employee or another private entrepreneur. In case the private

[9]This result has been supported by the work of Knack and Keefer (1995), Brunetti et al. (1998), Mauro (1998), Chong and Calderon (2000) and others. Furthermore, Ades and Di Tella (1997) present a formal model. Note that Mauro (1995) and other authors also find a significant effect of corruption on growth. For this direct effect, Ott (2000); Barreto (2000); Ehrlich and Lui (1999); Murphy et al. (1993); Li et al. (2000) and others provide formal models.

[10]See North (1990) and Husted (1999) for theoretical treatment of the topic and Tanzi and Davoodi (2000); Dollar and Kraay (2003); Kaufmann et al. (2000); Buscaglia (2001); Keefer and Knack (1997) for empirical evidence; among others.

entrepreneur plays against a corrupt government employee, the corrupt government employee siphons the surplus and the private entrepreneur is left without return. A fair government employee earns the government wage no matter whom he plays against. The corrupt government employee also earns the government wage if playing against another government employee. However, if he plays against a private entrepreneur, he extorts the surplus from private sector activity but bears the costs of the corrupt act - that is, the probability of getting caught red-handed and losing all his income in this case. The underlying game's payoffs represent the direct payoff flows of corruption. The surplus extorted by a corrupt government employee directly decreases the private entrepreneur's payoff. Thus the impact of corruption on a player's income is directly modelled in the underlying game.

In order to include the impact of income on the corruption rate we choose to relate the two by modelling institutional quality which is highly correlated with the corruption rate (e.g. Mauro, 1995). Our choice is guided by the fact that the relation between institutional quality and income is very well studied in economic research: The higher the income of a population, the better is the quality of its institutions.[11] We define institutional quality as the detection probability of corruption. If corrupt actions are detected with a high probability we say that institutional quality is high. Based on this, we can now model the impact of income on the corruption rate by assuming that the detection probability depends positively on income. Note that a high detection probability results in high costs of corruption. The costs of corruption determine the corrupt strategy's payoff and are therefore relevant for the decision to imitate it.

This completes the inclusion of the feedback effect in the model: On the one hand, income is a function of the strategy frequencies in the population and thus of the corruption rate. On the other hand, high income results in a high detection probability and decreases the incentive to corrupt.

This modelling idea requires an extension of the standard framework of an evolutionary game with replicator dynamics. It is necessary that the stage game depends on the strategy frequencies as explained above. We therefore define a new class of games which we call *frequency dependent evolutionary games*. We prove that frequency dependent evolutionary games dispose the fundamental properties of evolutionary games, such as uniqueness of solutions and invariance of the simplex and its boundaries. We refer to the frequency dependent evolutionary game featuring the effects described above as the corruption game. Accordingly we refer to the dynamics implied by the corruption game as the dynamics of corruption. As an equilibrium concept we adopt the evolutionary equilibria suggested by Friedman (1998). We calculate the evolutionary equilibria of the corruption game by applying local and global theory of nonlinear dynamical systems. In the Appendix we extend Liapunov's Theorem to critical points that lie on the simplex's boundary for the simplex invariant dynamics. This achievement is a most helpful result when analyzing frequency dependent evolutionary games.

[11] More income increases the value of property rights. "In sum, the basic structure of property rights is determined by the state and reflects the preferences and constraints of those who control the state" (Eggertsson, 1990, p. 79).

Our results are the following. We first allow for a general corruption detection probability function. A clean equilibrium where corruption is extinguished can only exist if detection probability is high in absence of corruption. As a consequence, populations with highly inefficient judicial institutions are not expected to free themselves from corruption.

In a corrupt equilibrium private activity is driven out of the game and only corrupt activity prevails. Such an equilibrium always exists if the detection probability is low in presence of a high corruption rate. We conclude that if corruption in the government affects the judicial institution's efficiency negatively, the possibility of a population being trapped in a corrupt equilibrium will prevail.

A hybrid equilibrium exists for a detection probability that is neither high nor low.

In the next step we specify a detection function that depends on income. The corruption game with such a detection function pays respect to the feedback effect between income and corruption rate. We find that a population can either converge to a clean equilibrium or to a corrupt equilibrium. It depends on the initial conditions to which equilibrium a population converges in the time limit. We analyze the respective basins of attraction. Our results suggest the existence of a pivot as shown by Paldam (2002) in an empirical study. The nature of the pivot is such that low income countries are prone to become trapped in the corrupt equilibrium for quite low corruption rates. In contrast, high income countries instead, seem to be more resistant to corruption in the sense that they converge to the clean equilibrium even when a moderately high initial corruption rate is present.

The plan of the present chapter is as follows. We first explain corruption in a standard evolutionary game and provide the intuition of the replicator dynamics by deriving them from an imitation rule. We analyze this basic corruption game and motivate the extension to implement a frequency dependent evolutionary game for the study of corruption. In Section 1.3.1 we define frequency dependent games and prove the fundamental results. In the subsequent section we analyze the corruption game by calculating the evolutionary equilibria and describing the global behavior of the solutions. Section 1.4 covers our conclusions. Note that for better readability all proofs of Propositions and Lemmas are collected in Appendix 1.A.

1.2 The Model

In Section 1.2.1 we define the *basic corruption game*, the evolutionary game which serves as our basic model. We present its stage game and explain why the assumed imitation rule generally leads to replicator dynamics. In addition, we introduce the notion of an evolutionary equilibrium as an equilibrium concept and discuss the evolutionary equilibria of the basic corruption game. In the following Section 1.2.2, we motivate the extension of standard evolutionary game theory to frequency dependent evolutionary games. We present the main ideas and refer to the related literature.

1.2.1 The Standard Framework

An evolutionary game describes strategic interaction over time. It is defined by the populations of players, a state space of strategies, a stage game, and an adaptation rule which determines the dynamic adjustment process.

The Population

The basic corruption game is a one-population game. We assume the population to consist of a continuum of infinitely-lived players. This assumption

has several well-discussed implications (see Friedman, 1998, for a complete list).

Firstly, the state space of strategies is continuous. In adherence to continuous time this allows us to specify the dynamics of a game as a system of ordinary differential equations.

Secondly, for an infinite number of players the law of large numbers can be invoked. This allows us to ignore random fluctuations and differing perceptions of the current state.

Lastly, an infinite number of players motivates the myopia assumption inherent to our dynamic adjustment process specified below. Players' influence on population are so small that players do not attempt to influence other players' future actions.

The Strategies

The basic corruption game consists of three strategies: strategy 1, strategy 2, and strategy 3. Hence, the pure-strategy set S of any player is $S = \{1, 2, 3\}$. To simplify interpretations, we presume that individuals only play pure strategies.[12] It is convenient for upcoming calculations to introduce the following notation. If an individual plays strategy i, $i \in S$, we denote his strategy choice σ_i as a vector in $\mathbb{R}^3$:

$$\sigma_i \in \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

representing a player's choice of strategy 1, 2, or 3 respectively.

The fraction of the population playing strategy i at time t is denoted by $x_i(t) \in [0, 1]$. The fractions of the population playing the three strategies (also called strategy frequencies) are the variables in our model. In our analysis we intend to describe the changes of these strategy frequency over time. By doing so, we are able to find the conditions under which corruption prevails or diminishes, and what the consequences for the payoffs of the players are.

[12] Note that this assumption is not necessary for the dynamic adjustment process we aim for. For example, Hofbauer and Sigmund (1998) discuss the replicator dynamics under the assumption that mixed strategies are played.

The *strategy state* of the game,

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix},$$

specifies the frequency of each of the three strategies in t. We will drop the time index t if there is no danger of misunderstanding.

The set of feasible strategy states is called the *strategy state space*. Since strategy frequencies must add up to one, the strategy state space for an evolutionary game is a simplex. We define the simplex of dimension $n-1$ as

$$\Sigma_{n-1} = \left\{ x(t) \in \mathbb{R}^n \,\middle|\, x_i(t) \geq 0 \text{ and } \sum_{i=1}^{n} x_i(t) = 1 \text{ for } i = 1, ..., n \right\}.$$

The strategy state space of the basic corruption game is Σ_2.

In the basic corruption game Strategy 1 represents the choice of holding down a job as government employee while not exploiting the power of the position. That is, a player choosing Strategy 1 acts as a *fair government employee*. Strategy 2 also comprises to serve in public service, but in contrast to Strategy 1, the player now abuses the power of the public role for private benefits. Strategy 2 can therefore be referred to as the strategy of a *corrupt government employee*. The third option is to pursue a private sector activity. We refer to Strategy 3 as the strategy of a *private entrepreneur*. Sometimes it is convenient to speak of public servants or government employees generally; in this case we refer to the total of players with Strategy 1 and 2. The share of public servants is abbreviated with $x_G = x_1 + x_2$. We also use the word administration for the government employees.

In our application, each pairwise encounter of two players is interpreted as one economic interaction. This means that each play of the stage game is considered as one economic act. The strategy choice of a player represents his decision which sector to direct his manpower to. The greater the share of individuals working in the public sector, x_G, the greater the share of economic activity that is processed entirely within the public sector (games played among government employees) or with the help of the public sector (games between private entrepreneurs and public servants).[13]

[13] The share of economic activity taking place within the government (administration) is $\Pr(x_G \geq y)^2 = x_G^2$ and the share of economic activity happening within the private sector is $\Pr(y \geq x_G)^2 = (1-x_G)^2 = x_3{}^2$.

Note that our choice of a one-population model has three implications for the interpretation of the basic corruption game. First, the size of government, measured as the share of agents employing strategy 1 or 2, is endogenous. Second, government employees play against themselves. These interactions can be interpreted as organizational and administrative tasks within a government which we believe to be a realistic feature in a model including the public sector.[14] Third, the private sector cannot elude the interaction with the public sector, no matter whether it is beneficial or damaging to its business.[15] In the next subsection we describe how the three strategies are presumed to interact.

The Stage Game

The stage game characterizes the strategic interaction of two players at any point in time. The stage game of the basic corruption game is a normal form game. It is defined by an expected payoff function $f(\sigma_i, x)$, where σ_i is the strategy choice of a specific player and $x \in \Sigma$ is the state of the game.

As in most of the existing literature, we adopt a linear expected payoff function originally employed in Maynard Smith (1982) and depict the game as a payoff matrix A. In every period players are drawn randomly and pairwise to play the stage game and receive the expected payoff $f(\sigma_i, x) = \sigma_i^T A x$.[16]

The assumptions for A are as follows. A fair government employee receives the wage w at any point in time, independently of his opponent. The interpretation is that he is paid w no matter whether he is busy mainly

[14] The rate at which the economic activities taking place within the administration increase with a marginal change of x_G is inherent in the game structure: $\frac{\partial \Pr(x_G \geq x)^2}{\partial x_G} = 2x_G$.

[15] Note that this feature is similar to a type of world Niskanen proposes: sponsors (in our case these are the private entrepreneurs as we will see in the next subsection) are passive in accepting the output proposal of bureaucracy without any careful monitoring or evaluation of alternatives (Niskanen, 1996).

[16] As common in the literature, we do not differentiate between the expected payoff against the population and the realized payoff of a specific stage game played. There are several reasons for that (Friedman, 1998): First, in large populations such as ours the expected payoff is a sufficient statistic. Second, payoffs are often not generated by random pairwise encounters, but by general interactions such as markets, and are therefore not stochastic.

delivering services to the private sector or doing work within the administration. The corrupt government employee, too, receives the wage w at any point in time, but additionally seizes a corruption income when playing against a private entrepreneur. The private entrepreneur generates a surplus s when being paired with another private entrepreneur or a fair government employee, and loses the fruits of his work when encountering a corrupt government employee. The corruption income of the corrupt public servant can be specified as $s - c$, where c depicts the individual costs of corruption. These assumptions lead to

$$A = \begin{pmatrix} w & w & w \\ w & w & w+s-c \\ s & 0 & s \end{pmatrix} .$$

It is a simplifying assumption that the payoff of a private entrepreneur does not vary between playing against a fair public servant and playing against a private entrepreneur. From our view this simplification is justified by the following interpretation: an entrepreneur makes the surplus s from his business activity while losing it with probability x_2, because he cannot circumvent interaction with the government.[17] Changing the payoff for Strategy 3 when playing against Strategy 1 (this is element a_{31} of matrix A), amounts to making a statement about the efficiency of public service. The reason is that x_1 would then affect the average payoff of Strategy 3. If $a_{31} = s$ though, then it is only x_2 that influences the private entrepreneur's average payoff. Although we have an idea in which direction we could change a_{31}, we do not want to include such an effect in our model. It is our aim to analyze the impact of corruption and we do not want to blur the results with other effects.

The Imitation Rule and the Adjustment Dynamics

Now let us describe in the following how the players select their strategies.

As mentioned above, our analysis is based upon the hypothesis, that strategy selection by imitation is a realistic assumption when describing illegal behavior. Players imitate the strategy of other, more successful players, where success refers to greater expected payoff. They base their decision

[17] Imagine that the entrepreneur needs to collect permits or certificates from the administration, has to pay taxes to the government, or is subjected to controls by law.

which strategy to imitate on sporadically and imperfectly observed expected payoffs and behavior. Information is scarce because knowledgeable players hide the information for which they can be prosecuted. To derive adjustment dynamics we slightly modify a model in Weibull (1995, p. 155). Note that time is continuous in our model.

Two elements are needed to find a general specification of the adjustment dynamics. With probability $\rho_i(x)$ players playing strategy i (called i-players in the following) review their strategy choice at any point in time. With probability $\varphi_i^j(x)$ a reviewing i-player switches to strategy j, $j \in S$, at any point in time. The share of i-players that will imitate another strategy is

$$\sum_{j \in S\setminus i} x_i \rho_i(x) \varphi_i^j(x) = x_i \rho_i(x) \sum_{j \in S\setminus i} \varphi_i^j(x) = x_i \rho_i(x)(1 - \varphi_i^i(x))$$
$$= x_i \rho_i(x) - x_i \rho_i(x) \phi_i^i(x) .$$

The share of players imitating i that have played a different strategy before is

$$\sum_{j \in S\setminus i} x_j \rho_j(x) \varphi_j^i(x) = \sum_{j \in S} x_j \rho_j(x) \varphi_j^i(x) - x_i \rho_i(x) \varphi_i^i(x) .$$

This leads to a net effect in the share of i-players of

$$\dot{x}_i = \sum_{j \in S} x_j \rho_j(x) \varphi_j^i(x) - x_i \rho_i(x) . \tag{1.1}$$

We can now further specify this general adjustment dynamics by making assumptions on $\rho_i(x)$ and $\varphi_j^i(x)$. First, we presume that players continuously review their strategies and set $\rho_i(x) = 1 \ \forall i$. Second, for specifying $\varphi_j^i(x)$ we suppose the following: At time t, every player samples an opponent with a probability equal for all opponents. The player observes the opponent's strategy and, with some noise, the opponent's expected payoff. Therewith, a j-player sampling an i-player observes $f(\sigma_i, x) - \varepsilon$ with probability x_i, where ε is a random variable with a continuously differentiable cumulative distribution function $\Phi : \mathbb{R} \rightarrow [0, 1]$.

We assume the following *imitation rule*: A j-player imitates strategy i when his own expected payoff (known without noise) is smaller than the observed expected payoff of strategy i. That is, he imitates strategy i if

$f(\sigma_j, x) < f(\sigma_i, x) - \varepsilon$. The probability that a j-player imitates i is $\Pr[\varepsilon < f(\sigma_i, x) - f(\sigma_j, x)] = \Phi[f(\sigma_i - \sigma_j, x)]$. From this we can conclude that

$$\begin{aligned} \varphi_j^i(x) &= x_i\,\Phi[f(\sigma_i - \sigma_j, x)] \quad \forall\, i \neq j\,, \\ \varphi_j^j(x) &= 1 - \sum_{i \in S\setminus j} x_i\,\Phi[f(\sigma_i - \sigma_j, x)]\,. \end{aligned}$$

Plugging this last expression and $\rho_i(x) = 1$ into (1.1), the adjustment dynamics takes the form

$$\begin{aligned} \dot{x}_i &= \sum_{j \in S} x_j \varphi_j^i(x) - x_i = \sum_{j \in S\setminus i} x_j \varphi_j^i(x) + x_i(\varphi_i^i(x) - 1) \\ &= \sum_{j \in S\setminus i} x_j\, x_i \Phi[f(\sigma_i - \sigma_j, x)] + x_i \left(- \sum_{j \in S\setminus i} x_j \Phi[f(\sigma_j - \sigma_i, x)] \right) \\ &= x_i \sum_{j \in S} x_j \Big(\Phi[f(\sigma_i - \sigma_j, x)] - \Phi[f(\sigma_j - \sigma_i, x)] \Big)\,. \end{aligned}$$

Finally we must specify the cumulative distribution function Φ. We assume a uniformly distributed error term ε over the interval of possible expected payoff differences. The function Φ is linear, $\Phi(y) = \alpha + \beta y$, and the adjustment dynamics derived from the imitation rule simplifies to

$$\begin{aligned} \dot{x}_i &= 2\beta x_i \sum_{j \in S} x_j f(\sigma_i - \sigma_j, x) = 2\beta x_i \left(\sum_{j \in S} x_j f(\sigma_i, x) - \sum_{j \in S} x_j f(\sigma_j, x) \right) \\ &= 2\beta x_i \left(f(\sigma_i, x) - f(x, x) \right)\,. \end{aligned}$$

Except from a time rescaling, our dynamics is thus equal to the replicator dynamics (see Taylor and Jonker, 1978; Schuster and Sigmund, 1983),

$$\dot{x}_i = x_i \left(f(\sigma_i, x) - f(x, x) \right) \quad \forall i \in S\,. \tag{1.2}$$

Since we do not focus on rates of convergence in this chapter, we can continue using equation (1.2). We alternatively will refer to it as the replicator dynamics or the imitation dynamics. Note that in our model, by equation (1.2), strategy selection from myopic imitation leads to a deterministic, continuous-time, continuous-state dynamics. Furthermore, note that equation (1.2) is a system of ordinary differential equations. For a simplified

notation, we define the system's right hand side as the (Lipschitz continuous) function $F : \Sigma \to \Sigma$, and can now write the imitation dynamics as $\dot{x} = F(x)$.[18]

Equilibrium Concept and Equilibria of the Basic Corruption Game

The most common equilibrium concept in biological literature concerned with evolutionary games is the ESS, which stands for Evolutionary Stable Strategy (see Hofbauer and Sigmund, 1988, for a thorough treatment). Orthodox game theory literature tends to focus on NE, the Nash-Equilibrium. As Friedman (1998) clearly points out, both the ESS and the NE, are *static* equilibrium concepts that rest upon the payoff function of the stage game. In Section 1.2.2 we abandon the assumption of a constant stage game wherefore a dynamic equilibrium concept needs to be employed. Instead of referring to a constant stage game, it must assure stability of F in an equilibrium. In Definition 1 we specify the *Evolutionary Equilibrium* EE in terms of the mathematical definitions of function F's stability in a critical point (also called stationary point or fixed points). The term evolutionary equilibrium was introduced by Hirshleifer (1982).

Definition 1 *A strategy state $x \in \Sigma_{n-1}$ is an evolutionary equilibrium of an evolutionary game if x is an attractor*[19] *of the dynamical system $\dot{x} = F(x)$ defining the game's adjustment dynamics.*

What is the interpretation of our equilibrium concept's definition for evolutionary games? As mentioned above, we are interested in how the imitation dynamics changes the strategy frequencies over time. By solving equation

[18] The replicator dynamics are simplex invariant:

$$\begin{aligned} \sum_{i \in S} \dot{x}_i &= \sum_{i \in S} x_i \left(f(\sigma_i, x) - f(x, x) \right) = \sum_{i \in S} x_i f(\sigma_i, x) - f(x, x) \sum_{i \in S} x_i \\ &= f(x, x) - f(x, x) = 0\,. \end{aligned}$$

[19] An attractor is defined as an asymptotically stable non-wandering set (in our case the only possible non-wandering sets are critical points and points on limit cycles or graphics). For definitions of critical points (also called equilibrium points or fixed points), limit cycles, graphics (also called separatrix cycles), asymptotic stability, and non-wandering sets see Perko (2000) or any textbook on dynamic systems.

(1.2), we can describe which strategies will be imitated more or less frequently for every strategy state in the state space. Thus, we can calculate which strategy frequencies a population exhibits over time after having been in a certain strategy state. Such a "solution path" for a given initial strategy state is called a solution trajectory.

An evolutionary equilibrium is a subset of state space Σ_{n-1} which a solution trajectory does not leave once reached. Additionally, if a solution trajectory of the dynamics starts sufficiently close to the evolutionary equilibrium, it remains close and converges asymptotically to the evolutionary equilibrium over time. The open set of points in Σ_{n-1} converging to a given EE are called its basin of attraction. We now present the evolutionary equilibria of the basic corruption game.

Proposition 1 *The strategy states in the set* $\{(x_1, 1-x_1, 0) \,|\, x_1 < \frac{w}{s}\}$ *are evolutionary equilibria of the basic corruption game. And if* $w < \min\{s, c\}$*, the strategy state* $x = (0, 0, 1)$ *is an evolutionary equilibrium additionally.*

Proposition 1 describes the evolutionary equilibria for all parameter values. The dynamics shows three different cases of combinations of evolutionary equilibria.

- If $s < w$ then $\{(x_1, 1-x_1, 0) \,|\, 0 \leq x_1 \leq 1\}$ is a set of EE (Case 1).
- If $w < s$ then $\{(x_1, 1-x_1, 0) \,|\, 0 \leq x_1 < \frac{w}{s}\}$ is a set of EE.
  - If $c < w$, then no other EE than the specified set exist (Case 2).
  - If $w < c$, then $x = (0, 0, 1)$ is an EE additionally (Case 3).

Figure 1.1 shows the three different cases of combinations of evolutionary equilibria. What are the conclusions that can be drawn from Proposition 1?

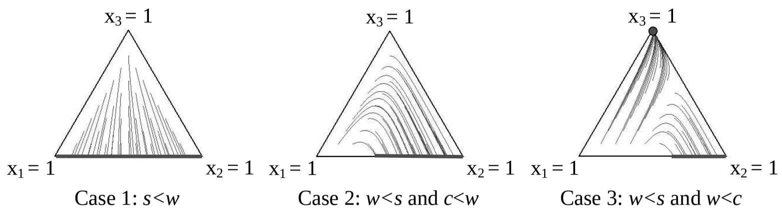


Figure 1.1: The three combinations of evolutionary equilibria.

If the government wage exceeds the surplus from private entrepreneurship, the private business will be crowded out by the strategies to engage with the government. In the opposite case, i.e. if the surplus from private activity is higher than government wage, the possibility that the population converges towards a state without private entrepreneurs persists. Only if punishment for corrupt behavior is sufficiently high ($c > w$) in combination with a government wage smaller than private surplus ($w < s$), a population can converges to a state with only private activity for some initial conditions. Note that there exists no evolutionary equilibrium in which both private and government activity occur.

The results we obtain by analyzing the basic corruption game are not plausible since we do not observe the described evolutionary equilibria in any country. For instance, in a real economy it is hardly imaginable that private activity is given up if it yields a high surplus compared to government salary and if the corruption rate is moderate. Furthermore we do not observe that government activity vanishes entirely if the individual costs of corruption are high. It is unsatisfactory that no evolutionary equilibria with positive frequencies for both private and government strategies exist. This motivates us to extend the model.

1.2.2 The Extension of the Standard Framework

As discussed above, the analysis of the basic corruption game is unsatisfactory. The reason is our assumption of a constant stage game. We do not expect a government salary to be constant when private sector activity is driven to zero. Similarly, we must take into account that the costs of corruption depend on how much resources are invested to reduce it: We argued above that these costs depend on the players' income.

In the next section, we will include these fundamental features in our analysis. The extended version of the basic corruption game is called *corruption game*. It is obvious that these extensions depart from standard evolutionary game theory since the stage game of the corruption game must now be dependent on the strategy state. We refer to this new class of games with strategy state dependent stage games as the class of *frequency dependent evolutionary games*. We will define it in the next section.

Very little research has been done on frequency dependent evolutionary games. The notion *game with frequency dependent payoffs* is due to Brenner and Witt (2003) discussed below. The concept itself dates back to Joosten et al. (1994), who first described games with changing payoffs, in a different context though. A few other examples of frequency dependent games (not evolutionary ones) are provided by Joosten et al. (2000) who introduce frequency dependent payoffs in the setup of stochastic games. They extend the concept of the Folk Theorem to capture the equilibria.

Brenner and Witt (2003) analyze evolutionary games with two-strategy, two-player stage games, where the dynamics is derived from individual strategy adaptation due to reinforcement learning in the variant of meliorating behavior. The resulting learning process changes the strategy choices in the population always in the same direction as the replicator dynamics. The authors describe the stationary state of a game that has a dominant strategy at every point in time. The players' payoffs depend on the frequency with which the dominant strategy has been played over all previous time periods.

Our analysis differs in several aspects from the setup of Brenner and Witt (2003). Firstly, we are concerned with the replicator dynamics. Secondly, our stage game is a three strategy game. Thirdly, we assume that the payoffs of the stage game in some period only depend on the frequencies of strategies played in that same period. Hence, strategy frequencies of prior periods do not enter the payoffs of the stage game. However these historical strategy frequencies have an effect on the stage game indirectly, namely through their impact on the present strategy frequencies.

1.3 Frequency Dependent Evolutionary Games

We first define the class of frequency dependent evolutionary games.

Definition 2 *An evolutionary game consisting of a population, a strategy state space, a strategy state dependent stage game, and a dynamic adjustment process, belongs to the class of frequency dependent evolutionary games.*

We adhere to the notation of Section 1.2.1 and only allude to the modifications necessary to transform an evolutionary game into a frequency dependent evolutionary game. The frequency dependent stage game of a frequency dependent evolutionary game is depicted by a matrix $\tilde{A}(x)$, where the matrix components $\tilde{a}_{kl}$ are functions of x. The corruption game thus has the payoff matrix

$$\tilde{A}(x) = \begin{pmatrix} \tilde{w}(x) & \tilde{w}(x) & \tilde{w}(x) \\ \tilde{w}(x) & \tilde{w}(x) & \tilde{w}(x)+\tilde{o}(x) \\ \tilde{s}(x) & 0 & \tilde{s}(x) \end{pmatrix}, \tag{1.3}$$

where the functions $\tilde{w}(x)$, $\tilde{s}(x)$, and $\tilde{o}(x)$ will be specified in Section 1.3.2. As a consequence of a frequency dependent stage game, the expected payoff of strategy choice σ_i is $f(\sigma_i, x) = \sigma_i^T A(x)x$, where f is not linear in x anymore. The replicator dynamics are given by (1.2) and can be written as

$$\dot{x}_i = x_i \left(\sigma^T A(x)x - xA(x)x\right) \quad \forall\, i \in S \tag{1.4}$$

for a frequency dependent evolutionary game.

For the rest of this section we proceed as follows. First, we prove in 1.3.1 that the general features of the replicator dynamics still hold for frequency dependent evolutionary games. Second, we deal with specific classes of frequency dependent evolutionary games. In 1.3.2 we analyze the corruption game which has the payoff matrix defined in (1.3). In Appendix 1.B we analyze the class of frequency dependent evolutionary games having stage games with only two pure strategies. We compare the evolutionary equilibria of the standard evolutionary games with two strategies with their frequency dependent counterparts.

1.3.1 Some General Results

The replicator dynamics of a frequency dependent evolutionary game is a system of differential equations as specified in (1.4). To guarantee that this system of differential equations induces a *well-defined* dynamics on the state space Σ_{n-1}, two conditions must be satisfied: (1) there is a unique solution for all initial conditions $x^0 \in \Sigma_{n-1}$, (2) these solutions must remain inside of Σ_{n-1}, i.e. $\phi_t(x^0) \in \Sigma_{n-1}$. Only if (1.4) satisfies these two fundamental requirements, the frequency dependent evolutionary game can be used as an economic model.

The following two propositions comprise the conditions for the system of differential equations (1.4) to be well-defined. Proposition 2 concerns the existence and the uniqueness of the solutions of system (1.4).

Proposition 2 *If all elements of $A(x)$ are Lipschitz-continuous functions, the replicator dynamics of a frequency dependent evolutionary game has a unique solution for every initial condition in the state space.*

Note that all proofs are in Appendix 1.A. The next proposition states that the unique solution of the replicator dynamics of a frequency dependent evolutionary lies in the interior of the game's state space.

Proposition 3 *If all elements of $A(x)$ are continuous functions, then the interior of simplex Σ and the boundary of the simplex Σ are both invariant under the replicator dynamics of a frequency dependent evolutionary game.*

Since Lipschitz-continuity implies continuity (see e.g. Walter, 1991), the differential equation system (1.4) induces a well-defined dynamics if we assume the elements of $A(x)$ to be Lipschitz-continuous functions of x.

Proposition 3 implies in particular

$$\sum_{i=1}^{N} x_i = 1 \quad \Rightarrow \quad \sum_{i=1}^{N} \dot{x}_i = 0 \quad \Rightarrow \quad \dot{x}_k = -\sum_{\substack{j=1 \\ j \neq k}}^{N} \dot{x}_j \,.$$

The change in the frequency of one strategy can be expressed through the changes in frequencies of the other strategies. This allows us to reduce the differential equation system of the replicator dynamics for frequency dependent evolutionary games by one equation.

A last property of the replicator dynamics, not as fundamental for the outcome as those established in Propositions 2 and 3, but very convenient for the calculations of the solution, is that it is invariant under positive continuous payoff transformations. Invariance under a positive continuous payoff transformation means that the functions in $A(x)$ can be multiplied by a positive real number without changing the solutions of the system. Similarly, adding or subtracting a continuous function from the columns of $A(x)$ does not change the replicator dynamics of a frequency dependent evolutionary game.

Proposition 4 *The replicator dynamics of a frequency dependent evolutionary game is invariant under positive continuous transformations of payoffs.*

In the Appendix 1.B we focus on the class of evolutionary games that have stage games with two strategies. We compare the evolutionary equilibria of the standard evolutionary games with those of the frequency dependent evolutionary games. To demonstrate the usefulness of our extension, we give an example that cannot be modelled satisfactorily by an evolutionary game with an invariable stage game.

In the next section, we apply Propositions 2, 3 and 4 to the corruption game in order to understand the dynamics of corruption.

1.3.2 The Corruption Game

In analyzing the corruption game we proceed as follows. Firstly, to assure that the model suits the aim of analyzing the dynamics of corruption, we make assumptions on the functions in the payoff matrix $\tilde{A}(x)$ of the corruption game. Then we present our main results. Proposition 5 describes the behavior of the solution trajectories that origin in the state space generally. We discuss the evolutionary equilibria given in the proposition and interpret their implications for the model. Finding a frequency dependent evolutionary game's EE can be a very hard task, since its differential equation system is highly nonlinear. Therefore we briefly describe the proof of Proposition 5 and allude to an extension of Liapunov's Theorem that is developed and proven in Appendix 1.A. This theorem can be very helpful for analyzing nonlinear dynamical systems with a simplex as a state space. Finally, we provide numerical solutions that provide further insight into the dynamics of corruption.

Defining the Corruption Game

The corruption game's payoff matrix $\tilde{A}(x)$ in (1.3) contains three functions: $\tilde{w}(x)$, the payoff resulting from government wage, $\tilde{o}(x)$, a corrupt government employee's expected additional payoff when encountering a private entrepreneur, and $\tilde{s}(x)$, the payoff from the return of private economic activity.

To specify these functions let us assume that government wage is a function of population income (and tax revenue therefore)[20] and that the government budget is financed through proportional taxes. Furthermore, let us assume that the government's budget is balanced at any point in time and that all legally earned payoffs are subject to taxes. The tax rate is denoted by τ. According to that we define $\tilde{w}(x) = (1-\tau)w(x)$, $\tilde{s}(x) = (1-\tau)s(x)$, and $\tilde{o}(x) = s(x) - c(x)$. Hence, we rewrite the payoff matrix as

$$A(x) = \begin{pmatrix} (1-\tau)w(x) & (1-\tau)w(x) & (1-\tau)w(x) \\ (1-\tau)w(x) & (1-\tau)w(x) & (1-\tau)w(x) + s(x) - c(x) \\ (1-\tau)s(x) & 0 & (1-\tau)s(x) \end{pmatrix}. \quad (1.5)$$

The function $c(x)$ depicts the expected costs of corrupting a private entrepreneur and absorbing his surplus $s(x)$ from private economic activity. The probability at which a corrupt activity is detected is denoted by $p(x)$, $p(x) \subseteq [0,1]$. We assume that if a corrupt government employee is caught red-handed, he is punished by having drawn off his net income, i.e.

$$c(x) \quad = \quad p(x)\left((1-\tau)w(x) + s(x)\right) .$$

Note that a corrupt government employee encountering a private entrepreneur obtains the payoff $w(x) + s(x)$ if $p(x) = 0$ and obtains a zero payoff if $p(x) = 1$. We leave detection probability $p(x)$ unspecified for the moment and describe the corruption game's dynamics for a general function $p(x)$. Later we will discuss the game with a specification of the function $p(x)$. For now we make the following assumption.

Assumption 1 *An increase in the share of corrupt government employees decreases the probability that a corrupt act is detected, i.e.,*

$$\frac{\partial p}{\partial x_2}(x) < 0 .$$

[20]The higher the GDP of a country, the higher is the government wage, though there exist differences in the percentage of government wage expenditure to GDP between countries. Hewitt and Van Rijckeghem (1995) find the following factors as determinants of government wage in a cross-country study: federal structure of political system, size of population in combination with per capita income, centralization of government, and government debt as factors. To keep our model simple, we abstract from these factors.

Further an increase in the share of corrupt government employees decreases the detection probability by more than an increase in the share of fair government employees, i.e.,

$$\frac{\partial p}{\partial x_2}(x) < \frac{\partial p}{\partial x_1}(x) .$$

Note that $\frac{\partial p}{\partial x_1}(x)$ can be positive or negative.

Though it is possible to analyze the game with $w(x)$, $s(x)$, and $c(x)$ being functions of x, we choose to normalize $s(x) = 1$. By that normalization we abandon the possibility of a frequency dependent surplus from private activity. Though we are aware of interesting specifications of $s(x)$,[21] we abstain from including such an effect. The reason is that we model the costs of corruption for the private sector in the payoff matrix directly, and do not want to lessen the explanatory power of the model by including further effects. Note that by setting $s(x) = 1$, government activities are of no direct utility for the private sector. But of course we can choose a $p(x)$ such that the share of fair government employees influences the costs of corruption.

These assumptions allow us to calculate the government wage explicitly. Tax revenue $r(x)$ is given by

$$\begin{aligned} r(x) &= \tau x^T \begin{pmatrix} w(x) & w(x) & w(x) \\ w(x) & w(x) & w(x) \\ 1 & 0 & 1 \end{pmatrix} x \\ &= \tau\left((x_1 + x_2)w(x) + (1 - x_1 - x_2)(1 - x_2)\right) . \end{aligned}$$

As mentioned above, the government is not able to tax the transfers from the private entrepreneur to the corrupt government employee. Thus, only legal payoffs are enlisted in the matrix used for tax revenue calculation. Now we can explicitly compute $w(x)$ using the budget constraint of the government:

$$\begin{aligned} w(x) &= \frac{r(x)}{x_1 + x_2}, \\ &= \frac{\tau\left((x_1 + x_2)w(x) + (1 - x_1 - x_2)(1 - x_2)\right)}{x_1 + x_2} . \end{aligned}$$

[21]Research in institutional economics suggests that s is higher the better the infrastructure of a country. Since we solely wish to analyze the impact of corruption, it is consequent to only have x_2 defining institutional quality.

By solving for government salary we receive

$$w(x) \quad = \quad \frac{\tau}{1-\tau}\frac{(1-x_1-x_2)(1-x_2)}{x_1+x_2} \, .$$

Plugging the explicit expression for $w(x)$ into the definition of $r(x)$ yields

$$r(x) \quad = \quad \frac{\tau}{1-\tau}(1-x_1-x_2)(1-x_2) \, .$$

This allows us to evaluate how the government wage and the tax revenue depend on x_1 and x_2.

$$\frac{\partial w(x)}{\partial x_2} < \frac{\partial w(x)}{\partial x_1} < 0 \quad \text{and} \quad \frac{\partial r(x)}{\partial x_2} < \frac{\partial r(x)}{\partial x_1} < 0 \, .$$

It is not surprising that the derivations of the government wage with respect to x_1 and x_2 are negative. Firstly, if either x_1 or x_2 increases, the frequency of private entrepreneurs decreases. However, this fraction of the population is solely responsible for the contribution of value to $r(x)$, since government employees only pay tax on their wages which are a fraction of $r(x)$; hence they cannot contribute to tax revenue positively. Secondly, the higher the number of government employees is, the lower is government salary for a given $r(x)$. The reason is our assumption of a balanced government budget: if x_1 or x_2 increases, $r(x)$ has to be split among more employees, so each gets a smaller wage. We also observe that the derivation of government wage with respect to x_2 is smaller than the one with respect to x_1. The rationale for this is the following. The more corrupt government employees there are, the higher is the share of games played among corrupt employees and private entrepreneurs. This implies that more of the entrepreneurs' surpluses flow outside the taxation system because they become illegal income from corruption. This reduces $r(x)$ and therewith $w(x)$. This is supported by empirical studies (e.g. Hwang, 2002; Hewitt and Van Rijckeghem, 1995).

The Main Result - The Dynamics of Corruption

In the last subsection we specified the functions in the payoff matrix of the corruption game, except for $p(x)$ which we want to treat generally. Proposition 4 states that we can rewrite the payoff matrix (1.3) as

$$A(x) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1-c(x) \\ (1-\tau)(1-w(x)) & -(1-\tau)w(x) & (1-\tau)(1-w(x)) \end{pmatrix} \qquad (1.6)$$

without having changed its dynamics. The replicator dynamics for frequency dependent evolutionary game yields a system of three differential equations for the corruption game. Following Proposition 3, we drop the third equation for $\dot{x}_3$ and substitute x_3 by $1-x_1-x_2$. This leaves us with the planar system

$$\begin{aligned} \dot{x}_1 &= x_1(1-x_1-x_2)\left[(1-\tau)(w(x)-1)-x_2(\tau-c(x))\right] \\ \dot{x}_2 &= x_2(1-x_1-x_2)\left[(1-\tau)w(x)+(1-x_2)(\tau-c(x))\right] \, . \end{aligned} \qquad (1.7)$$

We describe the dynamics of corruption in the following proposition.

Proposition 5 *The corruption game can only have the following three critical points as evolutionary equilibria:*

- $(\tau, 0)$ *is an EE if* $\frac{1}{2-\tau} < p(\tau, 0)$,
- $(0, 1)$ *is an EE if* $p(0,1) < \tau$,
- $(0, \bar{x}_2)$ *exists as a critical point if there exists an* $\bar{x}_2$ *satisfying* $p(0, \bar{x}_2) = \frac{\tau}{(1-\bar{x}_2)^2\tau+\bar{x}_2}$ *and is an EE if* $\tau > \bar{x}_2$ *and if* $-\frac{\partial c}{\partial x_2}(0, \bar{x}_2) < \frac{\tau}{\bar{x}_2^2}$.

There always exists at least one evolutionary equilibrium.

According to Proposition 5, the corruption game can have seven different combinations of evolutionary equilibria. We show these in Figure 1.2.

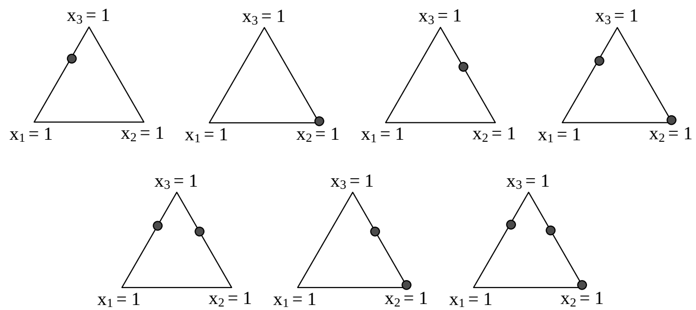


Figure 1.2: The possible combinations of evolutionary equilibria.

The proof of Proposition 5 is given in Appendix 1.A . In order to give an illustration how a frequency dependent evolutionary game has to be solved, let us present a verbal sketch of the proof. The proof is accomplished in two parts.

In Part I, we apply the local theory of nonlinear systems on (1.7). This involves calculating its critical points as well as determining the system's dynamical behavior in their vicinity. Critical points are found by setting the right hand side of (1.7) equal to zero and by solving for x_1 and x_2. There are several solutions and in many cases we find conditions for the existence of the critical point.

If a critical point is hyperbolic, we can determine its stability by applying the Hartman-Grobman Theorem. However, some critical points of system (1.7) are nonhyperbolic. Their asymptotic stability, or non-stability respectively, is typically more difficult to determine; one frequently used method is due to Liapunov. However, in frequency dependent evolutionary games, critical points may lie on the boundaries of the simplex where Liapunov's Theorem does not apply. Therefore, we prove an extension of Liapunov's Theorem to overcome the problems in determining the stability of nonhyperbolic critical points on simplex boundaries. Theorem 1 in Appendix 1.A states that the method due to Liapunov, applied as described in Theorem 1, can be used to determine the stability of a critical point on the simplex boundary as long as the simplex is invariant under the system's dynamics. As shown in Proposition 3, frequency dependent evolutionary games fulfill this necessary requirement under very general assumptions on the payoff functions. We believe that Theorem 1 can considerably simplify the analysis of frequency dependent evolutionary games in many cases.

In Part II, we check whether there exist other attracting sets of (1.7) in addition to the critical points found in Part I. We control for feasible attracting sets by applying Theorems of the global theory of nonlinear systems like the Poincaré-Bendixson Theorem and Theorems of Index Theory. We succeed in showing that the attracting sets of (1.7) must be critical points. Therefore we can conclude that we have found all evolutionary equilibria in Part I of the proof.

Interpreting the Main Result

We find three strategy states that can be evolutionary equilibria under our assumptions.

The critical point $(\tau, 0)$ always exists and is an evolutionary equilibrium if the detection probability of corrupt behavior is sufficiently high in $(\tau, 0)$. We refer to this EE as the *clean equilibrium* because all corrupt activity is crowded out by private activity and fair government service. If the function $p(x)$, a society's detection probability, takes high values for $x_2 \to 0$, we know a clean equilibrium exists. Therefore we conclude that for a society with a judiciary functioning efficiently at a low corruption rate, a clean equilibrium exists and at least some initial strategy states converge towards it. Note that a population may have a clean equilibrium for some tax rates, but not for others.

The second critical point, $(0, 1)$, also exists independently of the specification of $p(x)$. Furthermore, it is an evolutionary equilibrium if detection probability is low enough in $(0, 1)$. In this case we speak of a *corrupt equilibrium* because in this equilibrium all agents choose to be a corrupt government employee. Thus, if function $p(x)$ takes low values for high values of x_2, the corrupt equilibrium exists. A society's detection probability $p(x)$ takes low values for a high corruption rate if corruption badly affects a society's judiciary or its implementation. As soon as elements of a judiciary that are responsible for the detection of corruption can be corrupted, we expect that the society can be trapped in the corruption equilibrium for at least some initial strategy states.[22]

The last evolutionary equilibrium does not exist for all τ and all $p(x)$. If $(0, \bar{x}_2)$ exists and if the costs of corruption are not increasing too strongly

[22] The corrupt equilibrium may appear unlikely at first glance, because all players choose to be government employed although government wage converges to zero. We offer two arguments in favor of this evolutionary equilibrium. The first is an example: During Carlos Menem's last term of office as President of Argentina, 70% (!) of the labor force was employed by the local governments in many of the provinces (namely Chaco, Tucumán, La Rioja, and others). The majority of government employees barely worked yet collected their salary in the end of the month. At that time, Argentina certainly fulfilled the condition for the existence of the corrupt equilibrium: although the corruption rate was high, almost nobody was convicted for corrupt activities. Only 5% of Argentineans reported that they would seek judicial help in case of severe problems. The efficiency of the judiciary was too low for people to bother reporting corruption (and other crimes). The example is taken from TI's Daily Corruption News Service, http://www.transparency.org/cgi-bin/dcn-read.pl?citID=35148. The second argument is a theoretical one. The dynamics converge at an extremely low speed towards $(0, 1)$, so our model does not actually suggest any observations of $(0, 1)$ or its closest vicinity.

in x_2, then it is an EE of the corruption game. We refer to it as the *hybrid equilibrium* because the government service is entirely corrupt but does not suppress private economic activity. If a society's detection probability $p(x)$ does not depend strongly on the corruption rate and is neither very high nor very low, then we expect the hybrid equilibrium to exist. Consequently, a society whose judiciary is not too strong but also not too badly affected by corruption may have a hybrid equilibrium and can converge to it at least for some initial strategy states.

The stability conditions for the three critical points are independent of each other. Therefore the specification of $p(x)$ and the tax rate τ may imply any combination of the three evolutionary equilibria for the corruption game. By definition the EE are attractors of (1.7). Since Proposition 5 denotes *all* EE, every trajectory through an initial strategy state $x_0 \in \Sigma_2$ - except the separatrices of system (1.7) - converges to one of the EE. The separatrices of a dynamical system are those trajectories that approach a saddle point in the limit. The corruption game can have two saddle points at most; these are $(0, \bar{x}_2)$ and $(\tau - \hat{x}_2, \hat{x}_2)$ (see Proof of Proposition 5). So we have at most two separatrices within the simplex. If two or three EE exist for a given $p(x)$ and τ respectively, then the initial state is decisive for the EE the population converges to. We are interested in determining the sets of initial strategy states that converge to a certain evolutionary equilibrium. These sets are called the evolutionary equilibria's *basins of attraction.*

In order to describe each equilibrium's basin of attraction formally, the system would have to be solved explicitly. As in the case with many nonlinear differential equation systems, this is not possible. Instead, we provide an example of a function $p(x)$ and describe the global behavior of the corruption game by numerical simulations.

The Feedback Effect

In the following we discuss the dynamics of corruption for the specific function

$$p(x_1, x_2) \quad = \quad (1 - x_1 - x_2)(1 - x_2)\,.$$

This choice of $p(x)$ is motivated by the results of empirical studies. For example, Chong and Calderon (2000) find that corruption has an impact on growth, but also that growth influences the corruption rate. With our choice of $p(x)$ we pay respect to that fact insofar as income influences the corruption rate. The reasoning in our model is as follows. The higher the income of a population is, the more valuable are property rights. Assume now that the players have the possibility to protect their property rights, for instance by establishing a judiciary that is independent of the government. The better such an independent judiciary works, the higher is the detection probability of corrupt activities. We suggest that the efficiency of the judiciary depends on the amount of financial resources available. Hence, populations are disposed to provide more resources to protect property rights the richer they are. Therefore it is reasonable to assume that $p(x)$ depends positively on income.

In the corruption game, the legal[23] population income is

$$l(x_1, x_2) \quad = \quad \frac{(1 - x_1 - x_2)(1 - x_2)}{1 - \tau}\,. \tag{1.8}$$

For simplicity, we implement $p(x)$ as a linear function of $l(x)$. Since $p(x)$ is a probability it can only take values in $[0, 1]$. Consequently, we define it as $p(x) = (1 - \tau)l(x)$. Thus the probability of getting caught after committing a corrupt activity is proportional to the population income. Note that $\frac{\partial p}{\partial x_2} < \frac{\partial p}{\partial x_1} < 0$. An increasing share of corrupt as well as fair government employees decreases $p(x)$. The reason is that only private activities generate income where government employees are financed over taxes and do not contribute to population income.

The clean equilibrium exists for $p(\tau, 0) > \frac{1}{2-\tau}$, i.e. for $\tau < \frac{1}{2}(3 - \sqrt{5}) = 0.38$. The corrupt equilibrium exists for $p(0, 1) < \tau$, which is always satisfied because $p(0, 1) = 0$. The critical point is a saddle because condition (1.20) is satisfied for $\bar{x}_2$ evaluated at (1.13). Note that $\bar{x}_2$ increases in τ.

According to these results, the model predicts the following for a society in which the detection probability of corrupt activity depends positively on income: If tax burden is moderate, it either converges to the clean or the corrupt equilibrium, depending on the initial strategy state. If taxes are very high, it converges to the corrupt equilibrium.

We display the solution trajectories in Figure 1.3. The bottom left vertex of the simplex represents the strategy state $(1, 0)$ (only fair government

[23] We think that earnings from corrupt activities must not be included when constructing an indicator of property rights value.

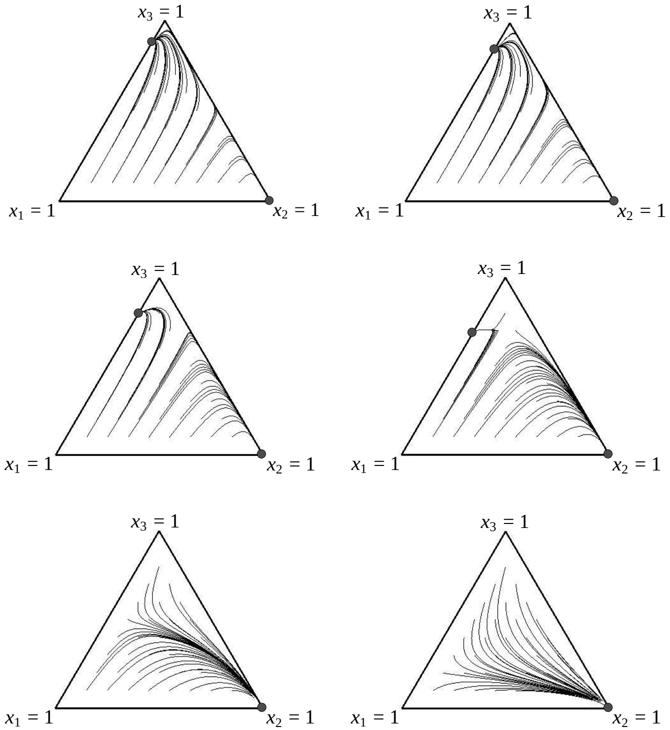


Figure 1.3: The dynamics of corruption with feedback effect.

employees), the bottom right vertex the strategy state $(0, 1)$ (only corrupt government employees), and the top vertex the strategy state $(0, 0)$ (only private entrepreneurs). The six pictures in Figure 1.3, are calculated with the tax rates $\tau = \{0.1; 0.15; 0.2; 0.3; 0.5; 0.8\}$. As mentioned above, the basin of attraction of the corrupt equilibrium broadens with τ increasing. The reason is that a high τ reduces the incentive of private activity because of a smaller relative payoff; firstly because more taxes have to be paid and secondly because government wage is higher.

How can we interpret the shape of the basins of attraction? The separatrix of point $(0, \bar{x}_2)$ separates the two basins of attraction. We find it by numerical simulation and plot it as a solid line in Figure 1.4 for a $\tau = 0.15$. The figure suggests that it is an approximately straight line parallel to

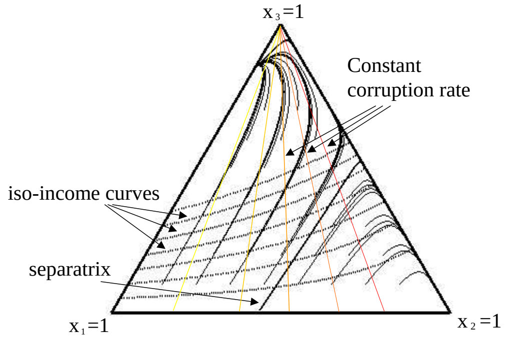


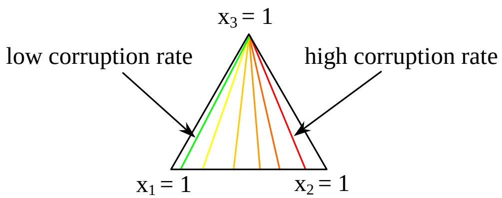
Figure 1.4: The basins of attraction and iso-income curves.

the edge of the simplex where $x_2 = 0$. All solution trajectories that start from an initial strategy state below this separatrix converge to the corrupt equilibrium. All others converge to the clean equilibrium. We discuss this observation in the following paragraph.

In empirical corruption research, the corruption rate is measured as the frequency of corrupt government employees within the government.[24] In our model such a measure coincides with the expression

$$\rho(x) \quad = \quad \frac{x_2}{x_1 + x_2}\,.$$

The corruption rate is constant on straight lines connecting the strategy state $(0, 0)$ with points on the simplex boundary where $x_3 = 0$:



[24] The most frequently used data is the CPI (corruption perception index) provided by TI (Transparency International), an international non-governmental organization devoted to combating corruption. The CPI "ranks countries in terms of the degree to which corruption is perceived to exist *among public officials and politicians* ", http://www.-transparency.org/cpi/index.html#cpi.

In Figure 1.4 we also depict *iso-income curves*. They are defined as sets of strategy states which generate the same legal income $l(x)$ defined in (1.8). The higher an iso-income curve lies in the simplex (the greater x_3), the greater is the income generated by its strategy states. The corrupt equilibrium generates an income of zero, $i(0,1) = 0$, the clean equilibrium generates an income of one, $l(\tau, 0) = 1$. We see that the separatrix crosses some of the iso-income curves. An intersection point that is more to the right of the simplex results in a higher income of the respective iso-income curve. We conclude that a population with a high income may start off with a high corruption rate and still converge to the clean equilibrium. Contrarily, a population with a low income that starts off with the same corruption rate may converge to the the corrupt equilibrium.

Paldam (2002) detects a similar effect empirically. In his paper, he first tests if cross-country data supports a model explaining corruption by cultural factors or a model containing economic factors. He finds that the economic model is superior and that GDP is the best predictor of the corruption rate. However, he also observes that countries either become too corrupt or too clean for their cultural affiliation. He proposes a seesaw dynamics to explain this observation. A so-called *pivot line* is defined as a corruption rate threshold. A corruption rate above the pivot line amplifies, a corruption rate below the pivot line dampens corruption. The data supports a pivot line such that most rich countries are above and most poor countries below. This observation is in accordance with the results of our model. We have shown analogously that the separatrix does not run along a constant corruption rate. The corrupt equilibrium's basin of attraction extends to lower corruption rates for small income populations than for high income populations.

Finally note that the shape and location of the separatrix depends on our exogenous variable τ. An enlargement of the basin of attraction of the clean equilibrium can be reached by lowering the tax rate. The empirical fact that government spending has a strong positive influence on corruption (Goel and Nelson, 1998) can therefore be seen as supportive of our model.

1.4 Conclusions

We analyze corruption in an evolutionary game where players choose between three strategies: to be a fair government employee, a corrupt government employee, or a private entrepreneur. Players are matched pairwise to play a stage game; each game played can be interpreted as an economic interaction. In such a model, the size of the government is endogenous. The stage game defines the following payoffs for the three strategies: Government employees have, regardless of their opponents, government wage as a payoff. However, the corrupt government employee additionally seizes the surplus of private economic activity and bears individual costs of corruption when playing against an entrepreneur. The private entrepreneur generates a surplus which he loses when playing against a corrupt entrepreneur. We assume that players observe the expected payoff of other opponents infrequently and with some noise. They imitate a strategy if they observe that it yields a higher expected payoff than their own strategy. Under certain assumptions this imitation rule leads to the replicator dynamics.

Although this evolutionary game generally captures the payoff transfers in a society where corrupt behavior interacts with private activity, this setup does not yield any plausible equilibria. The reason is that not all dynamical effects of corruption can be included in the standard setup.

Evolutionary games are defined by constant stage games which do not allow for feedback effects. Consequently, they are not suited to analyze economic situations in which population behavior affects the strategy payoffs of the stage game. However, there are many applications requiring this. Let us give an example to emphasize our point: People must choose to acquire a specific working skill in order to compete in the labor market. The expected salary is the higher, the less people possess a specific skill in the population. This situation can be modelled as a standard evolutionary game which perfectly describes individual decisions for different frequencies of skills. However, the standard setup does not allow to include more complex coherences like specialization effects. If many people possess the same skills, the probability of innovations rises which can have a major impact on the salaries paid in this sector.

Feedback effects are crucial for studying the dynamics of corruption: The incentives for private activity are smaller if income is lost to corrupt government employees. However, this is not the only effect of a high corruption rate. The reduced population income decreases government wage and the possibility to punish corrupt behavior, which increases the incentives for corrupt activity.

In order to comprehend feedback effects, we propose a new class of games, which we name the class of frequency dependent evolutionary games. We prove that frequency dependent evolutionary games fulfill the conditions for well defined dynamics under weak assumptions on the payoff functions. By analyzing the dynamics of corruption, we exemplarily solve a frequency dependent evolutionary game. The main difficulty of applying our new framework concerns the determination of the asymptotic stability of the game's equilibria. Therefore we present a theorem that extends Liapunov's Method to all possible equilibria of frequency dependent evolutionary games. Thereby we provide one approach that significantly reduces the technical difficulties of frequency dependent evolutionary games.

Let us now consider the frequency dependent evolutionary game describing corruption. It disposes an endogenous government wage and allows for the specification of institutional quality which can depend on population income. Our analysis of the corruption dynamics leads to the following conclusion: A society either converges to a clean equilibrium, a corrupt equilibrium, or a hybrid equilibrium.

In the clean equilibrium all players either choose to be a fair government employee or a private entrepreneur. It does only exist if the institutions punish corrupt behavior efficiently in case of a low corruption rate. Note that a society with weak institutions due to a high corruption rate can still converge to this equilibrium. The only condition for its existence is that institutions improve their efficiency with a decreasing corruption rate.

In the corrupt equilibrium all players choose to be corrupt government employees. This equilibrium exists whenever there is little punishment of corrupt behavior in case of a high corruption rate. Thus, if corrupt behavior decreases the individual costs of corruption or if institutions are generally weak, a society can converge to this equilibrium.

In the hybrid equilibrium all players either choose to be corrupt government employees or private entrepreneurs. This equilibrium exists if punishment of corruption is not strongly affected by the corruption rate and if it is neither severe nor weak.

If we include the feedback effects between population income and the corruption rate, i.e. if the quality of a society's institutions depends on the population income, the following results apply: The clean equilibrium only exists for tax rates below a certain threshold, the corrupt equilibrium does exist in any case. Let us assume that the tax rate is moderate enough for the clean equilibrium to exist. Our main finding then is the following: A society converges to the clean equilibrium if its corruption rate is below a certain threshold and it converges to the corrupt equilibrium if its corruption rate is above this threshold. The lower the population income is, the lower is this threshold. In other words, the lower the population income, the smaller must the corruption rate be for a society to converge to the clean equilibrium. High income populations however may start off with quite a high corruption rate but still converge to the clean equilibrium. A low income country with a high corruption rate is always trapped in the corrupt equilibrium.

1.A Appendix: Proofs

1.A.1 Proof of Proposition 1

The replicator dynamics (1.2) are invariant under positive affine transformations of payoffs (see e.g. Weibull, 1995, p. 73), hence we can redefine A as

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & s-c \\ s-w & -w & s-w \end{pmatrix}. \tag{1.9}$$

In order to find the critical points of the system defined by $\dot{x} = F(x)$, we have to solve

$$F(x) = \begin{pmatrix} x_1\left(e_1{}^T Ax - xAx\right) \\ x_2\left(e_2{}^T Ax - xAx\right) \\ x_3\left(e_3{}^T Ax - xAx\right) \end{pmatrix} = 0.$$

Due to the fact that the replicator dynamics is simplex invariant, we can drop one equation and replace one variable. We choose to drop the equation for $\dot{x}_3$, and eliminate x_3 by $1 - x_1 - x_2$. This leaves us with the reduced system

$$\begin{aligned} x_1(1-x_1-x_2)(w+x_2c-s) &= 0 \\ x_2(1-x_1-x_2)(w+x_2c-c) &= 0. \end{aligned}$$

We redefine $F : \mathbb{R}^2 \to \mathbb{R}^2$ as the left hand side of the reduced system. The function F is differentiable on $\mathbb{R}^2$ and therefore Lipschitz continuous on $\mathbb{R}^2$. Thus the Fundamental Existence-Uniqueness Theorem (Picard-Lindelöf) applies (Perko, 2000, p. 74). The critical points are listed in the first column of Table 1.1. Note that we assume that $s \neq w$ and $s \neq c$. The Jacobian $DF(x)$ is

$$\begin{pmatrix} (1-2x_1-x_2)(w+x_2c-s) & x_1(s-w+c-c(x_1-2x_2)) \\ x_2(c-w-x_2c) & (1-x_1-2x_2)(w-c+2x_2c)+x_2^2c \end{pmatrix}.$$

The critical points $(0,0)$ and $(0, \frac{w-c}{c})$ are hyperbolic because both $DF(0,0)$ and $DF(0, \frac{w-c}{c})$ have non-zero eigenvalues. We can therefore apply the

Hartman-Grobman Theorem (Perko, 2000) and derive the conditions for asymptotic stability from these eigenvalues. We find these conditions for asymptotic stability in the forth column of Table 1.1. It is obvious that $(0, \frac{c-w}{c})$ cannot be asymptotically stable.

Let us now tackle the stability of the nonhyperbolic critical points $(x_1, 1 - x_1)$. We know that the non-zero eigenvalue has to be negative for a critical point to be asymptotically stable (see Perko, 2000, Theorem 2, p. 130), so only $x \in \Sigma_2$, for which $w > x_1 s$ is true, can be asymptotically stable points.

Table 1.1 summarizes the results of this proof.

Table 1.1: Critical points of the basic corruption game.

Critical Point	Conditions to parameters for existence	Eigenvalues of $J_F(x_1, x_2)$	Asymptotic stability if
$x_1 = 0, x_2 = 0$	none	$w-s, w-c$	$w < \min\{s, c\}$
$x_1 = 0, x_2 = \frac{c-w}{c}$	$w \in (0, c]$	$\frac{w(c-s)}{c}, \frac{w(c-w)}{c}$	$c < \min\{s, w\}$
$x_1 = x_1, x_2 = 1 - x_1$	none	$0, x_1 s - w$	$w > x_1 s$

1.A.2 Proof of Proposition 2

Proposition 2 follows directly from The Fundamental Existence-Uniqueness Theorem (Picard-Lindelöf), Perko (e.g. 2000).

1.A.3 Proof of Proposition 3

For simplex invariancy of the replicator dynamics of a frequency dependent evolutionary game the following three conditions must be satisfied:

$$\sum_{i=1}^{N} \dot{x}_i = 0 \tag{1.10}$$

$$\lim_{x_i \to 0^+} \dot{x}_i = 0 \tag{1.11}$$

$$\lim_{x_i \to 1^-} \dot{x}_i \;=\; 0 \tag{1.12}$$

Condition (1.10) guarantees that the solution of the system satisfies $\sum_{i=1}^{N} x_i = 1$ if the initial condition is an element of simplex Σ. Conditions (1.11) and (1.12) impose the upper bound 1 and the lower bound 0 on the solution $x(t)$. The three together limit the solution $x_i(t)$ to the simplex Σ.

We introduce the following notation for row i of matrix A:

$$A_i \;=\; (a_{i1}, a_{i2}, ..., a_{i\,n-1}, a_{i\,n})\,.$$

Condition (1.11) can be written as

$$\lim_{x_i \to 0^+} \dot{x}_i = \lim_{x_i \to 0^+} x_i \left((A_i x) - \left(\sum_{j=1}^{N} x_j (A_j x) \right) \right) = \lim_{x_i \to 0^+} x_i g(x)\,,$$

where we denote the function in brackets by $g(x)$. If all elements of A are continuous functions on simplex Σ, then $g(x)$ is a continuous function on Σ because sums and products of continuous functions are continuous functions. The simplex Σ is compact, from this it follows (Theorem of Weierstrass) that $g(x)$ is compact (and therefore bounded). So we have

$$\lim_{x_i \to 0^+} \dot{x}_i = 0\,.$$

We next consider condition (1.12),

$$\lim_{x_i \to 1^-} \dot{x}_i \;=\; \lim_{x_i \to 1^-} x_i \left((A_i x) - \left(\sum_{j=1}^{N} x_j (A_j x) \right) \right)\,,$$

under the assumption that $a_{ij}(x)$ are continuous functions. By the reasoning above we know that sums and products of the functions $a_{ij}(x)$ are bounded and that limits on Σ are finite therefore. Thus we can write

$$\begin{aligned}
\lim_{x_i \to 1^-} \dot{x}_i &= \lim_{x_i \to 1^-} x_i (A_i x) - \lim_{x_i \to 1^-} \sum_{j=1}^{N} x_j (A_j x) \\
&= \lim_{x_i \to 1^-} (A_i x) - \sum_{j=1}^{N} \lim_{x_i \to 1^-} x_j (A_j x) \\
&= -\sum_{j \neq i} \lim_{x_i \to 1^-} x_j (A_j x)
\end{aligned}$$

For $x_i \to 1^-$, we have that $x_j \to 0^+$ for $j \neq i$. Again continuity of the functions $a_{ij}(x)$ on Σ implies that

$$\lim_{x_j \to 0^+} x_j (A_j x) = 0$$

and therefore we have

$$\lim_{x_i \to 1^-} \dot{x}_i \;=\; 0\,.$$

Finally condition (1.10) can be shown by summation of all equations in (1.4):

$$\begin{aligned}
\sum_{i=1}^{N} \dot{x}_i &= \sum_{i=1}^{N} x_i \left(e_i' A(x) x - x' A(x) x \right) \\
&= \sum_{i=1}^{N} x_i (e_i' A(x) x) - \sum_{i=1}^{N} x_i (x' A(x) x) \\
&= \sum_{i=1}^{N} (x_i e_i') A(x) x - (x' A(x) x) \sum_{i=1}^{N} x_i \\
&= \left(\sum_{i=1}^{N} x_i e_i' \right) A(x) x - x' A(x) x \\
&= x' A(x) x - x' A(x) x \;=\; 0\,.
\end{aligned}$$

From the above it is clear that if $x_i = 0$, we have $\dot{x}_i = 0$. Thus the boundary of Σ is invariant. When rewriting equation (1.4) as

$$\frac{\dot{x}_i}{x_i} \;=\; \left(\sigma^T A(x) x - x A(x) x \right) \quad \forall\, i \in S$$

we see that the differential equation system intuitively describes the relative change of the solutions $x_i(t)$. From this it is obvious that if $x_i(0) > 0 \Leftrightarrow x_i(t) > 0$. So the interior of Σ is invariant, too.

1.A.4 Proof of Proposition 4

If we multiply all payoffs with $\lambda > 0$, we can write the replicator dynamics as

$$\dot{x}_i \;=\; x_i \left(\sigma^T \lambda A(x) x - x \lambda A(x) x \right) \;=\; \lambda x_i \left(\sigma^T A(x) x - x A(x) x \right)\,,$$

which is a system with the same solutions $x_i(t)$ as (1.4).

Now let $B(x)$ be a matrix with n identical rows, which we denote by

$$b(x) = (b_1(x), b_2(x), ..., b_{n-1}(x), b_n(x))\,.$$

The elements of $B(x)$ are continuous functions on Σ.

$$\begin{aligned}\dot{x}_i &= x_i\left(\sigma^T\left[A(x)+B(x)\right]x - x\left[A(x)+B(x)\right]x\right)\\ &= x_i\left(\sigma^T A(x)x + b(x)x - xA(x)x - xb(x)x\right)\\ &= x_i\left(\sigma^T A(x)x - xA(x)x\right)\,.\end{aligned}$$

1.A.5 Proof of Proposition 5

We accomplish this proof in two parts. First, we apply the local theory of nonlinear systems and second, we use theorems of global theory of nonlinear systems to show that we have found all attractors of (1.7).

Local Theory of Nonlinear Systems

From Proposition 2 and the differentiability (and therefore Lipschitz-continuity) of the functions $w(x)$ and $c(x)$ we know that the differential equation system (1.7) has a unique solution. We redefine $F : \mathbb{R} \to \mathbb{R}$ as the right hand side of system (1.7).

$$\begin{aligned}F(x) &= \begin{pmatrix} F_1(x)\\ F_2(x)\end{pmatrix}\\ &= \begin{pmatrix} x_1(1-x_1-x_2)\left[(1-\tau)(w(x)-1) - x_2(\tau 1 - c(x))\right]\\ x_2(1-x_1-x_2)\left[(1-\tau)w(x) + (1-x_2)(\tau 1 - c(x))\right]\end{pmatrix}\,.\end{aligned}$$

In order to find the critical points of the system, we set $F(x) = 0$:

$$\begin{aligned}x_1(1-x_1-x_2)\left[(1-\tau)(w(x)-1) - x_2(\tau 1 - c(x))\right] &= 0\\ x_2(1-x_1-x_2)\left[(1-\tau)w(x) + (1-x_2)(\tau 1 - c(x))\right] &= 0\,.\end{aligned}$$

There are nine possibilities which are solution candidates for this equation system.

1) $x_1 = 0 \;\wedge\; x_2 = 0$

2) $x_1 = 0 \;\wedge\; (1-x_1-x_2) = 0$

3) $x_1 = 0 \;\wedge\; (1-\tau)w(x) + (1-x_2)(\tau - c(x)) = 0$

4) $(1-x_1-x_2) = 0 \;\wedge\; x_2 = 0$

5) $(1-x_1-x_2) = 0 \;\wedge\; (1-x_1-x_2) = 0$

6) $(1-x_1-x_2) = 0 \;\wedge\; (1-\tau)w(x) + (1-x_2)(\tau - c(x)) = 0$

7) $(1-\tau)(w(x)-1) - x_2(\tau - c(x)) = 0 \;\wedge\; x_2 = 0$

8) $(1-\tau)(w(x)-1) - x_2(\tau - c(x)) = 0 \;\wedge\; (1-x_1-x_2) = 0$

9) $(1-\tau)(w(x)-1) - x_2(\tau - c(x)) = 0$
$\wedge\; (1-\tau)w(x) + (1-x_2)(\tau - c(x)) = 0\,.$

Conditions 1), 2), and 4) state that the vertices of the simplex are fixed points.

Condition 5) gives us the edge of the simplex where $x_3 = 0$ as a set of critical points, conditions 6) and 8) give two single points in this set as fixed points.

Condition 3) gives us a critical point on the edge of the simplex where $x_1 = 0$. It only exists if there is a solution $\bar{x}_2$ which satisfies

$$(1-\tau)w(0,\bar{x}_2) + (1-\bar{x}_2)(\tau - c(0,\bar{x}_2)) = 0\,. \tag{1.13}$$

By plugging

$$w(0,\bar{x}_2) = \frac{\tau}{1-\tau}\frac{(1-\bar{x}_2)^2}{\bar{x}_2}$$

into (1.13), we receive

$$c(0,\bar{x}_2) = \frac{\tau}{\bar{x}_2}\,.$$

We substitute the expression for $c(0,\bar{x}_2)$ given by (1.15) into the last equation and get

$$p(0,\bar{x}_2) = \frac{\tau}{\tau(1-\bar{x}_2)^2 + \bar{x}_2}\,.$$

So the potential critical point $(0,\bar{x}_2)$ only exists if $p(0,x_2)$ intersects with the function $g(x_2) = \frac{\tau}{\tau(1-x_2)^2+x_2}$ at least once on $]0,1[$ for a given τ. The function $g(x_2)$ is strictly increasing in the parameter τ and can take values that are greater than one. We conclude that for high τ, an $\bar{x}_2$ will not exist. The function $g(x_2)$ is decreasing in x_2 when its image is in the interval $]0,1[$.

Condition 7) implies

$$\begin{aligned}(1-\tau)(w(x_1,0)-1) &= 0\\ w(x_1,0) &= 1\\ \frac{\tau}{1-\tau}\frac{1-x_1}{x_1} &= 1 \quad\Rightarrow\quad x_1=\tau\,,\end{aligned}$$

so $(\tau,0)$ is a fixed point.

Finally, condition 9) supports a critical point if there is a solution to

$$\begin{aligned}(1-\tau)(w(x_1,x_2)-1)-x_2(\tau-c(x_1,x_2)) &= 0\\ (1-\tau)w(x_1,x_2)+(1-x_2)(\tau-c(x_1,x_2)) &= 0\,.\end{aligned}$$

We rewrite the system as

$$\begin{aligned}(1-\tau)w(x_1,x_2)-(1-\tau)-x_2(\tau-c(x_1,x_2)) &= 0\\ (1-\tau)w(x_1,x_2)+(\tau-c(x_1,x_2))-x_2(\tau-c(x_1,x_2)) &= 0\end{aligned}$$

and find

$$c(x_1,x_2) = 1$$

by subtracting the two equations. Plugging this result into the first of the equations of the system gives

$$\begin{aligned}(1-\tau)(w(x_1,x_2)-1)-x_2(\tau-1) &= 0\\ (w(x_1,x_2)-1)+x_2 &= 0 \quad\rightarrow\quad w(x_1,x_2)=1-x_2\,.\end{aligned}$$

We now use the explicit expression for $w(x_1,x_2)$ and find

$$\begin{aligned}w(x_1,x_2)=1-x_2 \quad\rightarrow\quad \frac{\tau}{1-\tau}\frac{(1-x_1-x_2)(1-x_2)}{x_1+x_2} &= 1-x_2\\ \frac{\tau}{1-\tau}(1-x_1-x_2) &= x_1+x_2\\ \tau &= x_1+x_2\,.\end{aligned}$$

Let us denote $\hat{x}_2$ as the solution of $c(\tau-\hat{x}_2,\hat{x}_2)=1$, and we derive further

$$\begin{aligned}c(\tau-\hat{x}_2,\hat{x}_2) &= 1\\ p(\tau-\hat{x}_2,\hat{x}_2)((1-\tau)(1-\hat{x}_2)+1) &= 1\\ p(\tau-\hat{x}_2,\hat{x}_2) &= \frac{1}{(1-\tau)(1-\hat{x}_2)+1}\,.\end{aligned}$$

So, only if we assume a $p(x_1,x_2)$ such that $p(\tau-\hat{x}_2,\hat{x}_2)$ intersects with a function $h(x_2)$,

$$h(x_2) = \frac{1}{(1-\tau)(1-x_2)+1}\,,$$

on $[0,\tau]$, there exists a fixed point $(\tau-\hat{x}_2,\hat{x}_2)$. Note that $h(x_2)$ is increasing in τ and increasing in x_2. The image of the function $h(x_1)$ is in $[\frac{1}{2},1]$.

We enlist the critical points we have found in the first column of Table 1.2.

In a next step we want to analyze the stability of the critical points we have found above.

If a critical point x_0 is hyperbolic,[25] it is either a sink,[26] a saddle,[27] or a source[28] (Definitions e.g. by Perko, 2000, p. 102). It follows from the Hartman-Grobman Theorem[29] that sinks of a differential equation system are asymptotically stable and sources and saddles are unstable. So in order to determine if a hyperbolic critical point is asymptotically stable (and thus an EE) or not we only need to calculate the eigenvalues of the Jacobian of $F(x)$ evaluated at the critical point. Therefore, we first calculate the elements of the Jacobian $DF(x_1,x_2)$.

$$\frac{\partial F_1}{\partial x_1} = (1-2x_1-x_2)\Big((1-\tau)(w(x)-1)-x_2(\tau-c(x))\Big)$$

[25] None of the eigenvalues of $DF(x_0)$ has a zero real part.

[26] All eigenvalues of $DF(x_0)$ have negative real parts.

[27] All eigenvalues of $DF(x_0)$ have positive real parts.

[28] At least one eigenvalue of $DF(x_0)$ has a positive and at least one has a negative real part.

[29] The Hartman-Grobman Theorem states that if F is differentiable then there exists a homeomorphism that maps the trajectories in an open set around a hyperbolic critical point x_0 onto trajectories near x_0 of the linear system $\dot{x}=Ax$ with $A=DF(x_0)$. That is to say that near a hyperbolic critical point x_0 the nonlinear system $\dot{x}=F(x)$ has the same qualitative structure as the linear system $\dot{x}=Ax$ with $A=DF(x_0)$.

$$
\begin{aligned}
&\quad + x_1(1-x_1-x_2)\left((1-\tau)\frac{\partial w(x)}{\partial x_1} + x_2\frac{\partial c(x)}{\partial x_1}\right)\\
\frac{\partial F_1}{\partial x_2} &= -x_1\Big((1-\tau)(w(x)-1) - x_2(\tau - c(x))\Big)\\
&\quad + x_1(1-x_1-x_2)\left((1-\tau)\frac{\partial w(x)}{\partial x_2} - (\tau - c(x)) + x_2\frac{\partial c(x)}{\partial x_2}\right)\\
\frac{\partial F_2}{\partial x_1} &= -x_2\Big((1-\tau)w(x) + (1-x_2)(\tau - c(x))\Big)\\
&\quad + x_2(1-x_1-x_2)\left((1-\tau)\frac{\partial w(x)}{\partial x_1} - (1-x_2)\frac{\partial c(x)}{\partial x_1}\right)\\
\frac{\partial F_2}{\partial x_2} &= (1-x_1-2x_2)\Big((1-\tau)w(x) + (1-x_2)(\tau - c(x))\Big)\\
&\quad + x_2(1-x_1-x_2)\left((1-\tau)\frac{\partial w(x)}{\partial x_2} - (\tau - c) - (1-x_2)\frac{\partial c(x)}{\partial x_2}\right).
\end{aligned}
$$

For convenience we also restate

$$
\begin{aligned}
w(x) &= \frac{\tau}{1-\tau}\frac{(1-x_1-x_2)(1-x_2)}{x_1+x_2} \quad \text{and} && (1.14)\\
c(x) &= p(x)\Big((1-\tau)w(x)+1\Big) \quad \text{with} && (1.15)\\
\frac{\partial w}{\partial x_1} &= -\frac{\tau}{1-\tau}\left(\frac{1-x_2}{(x_1+x_2)^2}\right) && (1.16)\\
\frac{\partial w}{\partial x_2} &= \frac{\tau}{1-\tau}\left(1-\frac{1+x_1}{(x_1+x_2)^2}\right). && (1.17)
\end{aligned}
$$

We treat the critical points one by one in the order of Table 1.2.

Critical point $(0,0)$: It is easy to see that $\frac{\partial F_1}{\partial x_2}(0,0)=0$ and $\frac{\partial F_2}{\partial x_1}(0,0)=0$. The eigenvalues of $DF(0,0)$ are thus $\frac{\partial F_1}{\partial x_1}(0,0)$ and $\frac{\partial F_2}{\partial x_2}(0,0)$.[30] We find

$$
\begin{aligned}
\frac{\partial F_1}{\partial x_1}(0,0) &= (1-\tau)w(0,0)\\
\frac{\partial F_2}{\partial x_2}(0,0) &= (1-\tau)w(0,0)(1-p(0,0)) + \tau - p(0,0).
\end{aligned}
$$

The expression $(1-\tau)w(0,0)$ is clearly positive, because of

$$
\lim_{\substack{x_1\to 0\\ x_2\to 0}} w(x_1,x_2) = \infty.
$$

[30] The eigenvalues of a matrix $\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ are equal to a and d if either $b=0$ or $c=0$ or both.

If $p(0,0)$ such that $\frac{\partial F_2}{\partial x_2}(0,0)\neq 0$, then the critical point $(0,0)$ is hyperbolic and one positive eigenvalue is enough to know that it is unstable (see e.g. Perko, 2000, Theorem 2, p. 130). If $p(0,0)$ such that $\frac{\partial F_2}{\partial x_2}(0,0)=0$, then the second eigenvalue is equal to zero and $(0,0)$ is not hyperbolic. It is then an unstable node, a saddle, or a saddle-node and unstable therefore (see e.g. Perko, 2000, Theorem 1, p. 151).

Critical point $(0,1)$: The critical point $(0,1)$ does always exist and is not hyperbolic. In fact, we even have $DF(0,1)=0$, which indicates a very complex behavior of the system near the critical point. For most other critical points it is more convenient to analyze the behavior near them in $\mathbb{R}^2$ (although we are only interested in the dynamics on the simplex).[31] For showing asymptotic stability of $(0,1)$ however, we will restrict our analysis to the simplex, which makes our efforts more comprehensible in this case.

One method to show stability for critical points that are not hyperbolic is due to Liapunov. The theorem (see e.g. Perko, 2000, p. 131, Theorem 3) that states under which conditions the existence of a Liapunov function (defined below) implies (asymptotic) stability of a critical point only applies to critical points that are interior points of the definition space of $F(x)$. The critical point $(0,1)$ is a boundary point of the simplex though. Our proceeding is as follows. First, we will prove a new theorem that states that the existence of a Liapunov function guarantees asymptotic stability for a boundary point of the simplex if the simplex is invariant under $\dot{x}=F(x)$. Second, we will give an example of a Liapunov function for the system of the corruption game and show under which circumstances $(0,1)$ is asymptotically stable.

Theorem 1 *Let E be an open subset of Σ and $x_0\in\overline{E}$.*[32] *Suppose that $F(x)\in C^1(\overline{E})$ and $F(x_0)=0$, where the simplex is invariant under $\dot{x}=F(x)$. Suppose further that there exists a real valued function $V\in C^1(\overline{E})$*

[31] One reason is that the Hartman-Grobman Theorem requires open subsets containing the hyperbolic critical points. Most of our critical points are on the boundary of the simplex, however.

[32] Notation: $\overline{E}$ is the set of osculation points of E. Point x is an osculation point of E if $E\cap U_\rho(x)\neq\emptyset, \forall\rho\in\mathbb{R}_+$. The set $U_\rho(x)$ is the ρ-neighborhood of x (or the open sphere around x). It is defined as $U_\rho(x)=\{y\in\mathbb{R}^n; |x-y|<\rho\}$.

satisfying $V(x_0) = 0$ *and* $V(x) > 0$, $\forall x \in \overline{E} \setminus x_0$. *If* $\dot{V}(x) < 0\ \forall x \in E$, x_0 *is asymptotically stable.*

Proof of Theorem 1.[33] Function $V(x)$ is called a Liapunov function. We define $\phi_t(x)$ as the flow of system $\dot{x} = F(x)$. We can write

$$\dot{V}(x) = \frac{d}{dt} V(\phi_t(x))|_{t=0} = DV(x)F(x)\,. \tag{1.18}$$

The first equation is due to the definition of the flow of a differential equation system, the second equation is due to the chain rule.

Choose $\varepsilon > 0$ sufficiently small that $\overline{N_\varepsilon(x_0)} = \overline{U_\varepsilon(x_0) \cap \Sigma_2} \subset \overline{E}$. We define the compact set S_ε,

$$S_\varepsilon \;=\; \{x \in \mathbb{R}^2 |\ |x - x_0| = \varepsilon\} \cap \overline{N_\varepsilon(x_0)}\,.$$

Since $V(x)$ is continuous there exists a minimum m_ε of $V(x)$ on S_ε and $V(x) > 0$ for $x \in \overline{E} \setminus x_0$ implies $m_\varepsilon > 0$. We also have $V(x_0) = 0$ and since $V(x)$ is continuous there exists a δ such that $|x - x_0| < \delta$ implies $V(x) < m_\varepsilon$. Equations (1.18) imply that if $\dot{V}(x) < 0$ for $x \in E$, $V(x)$ is strictly decreasing along the trajectories of $\dot{x} = F(x)$. It follows that for all $\tilde{x} \in \overline{N_\delta(x_0)} = \overline{U_\delta(x_0) \cap \Sigma} \subset \overline{E}$ and $t > 0$ we have

$$V(\phi_t(\tilde{x})) < V(\tilde{x}) < m_\varepsilon\,. \tag{1.19}$$

Now suppose that for $\tilde{x}$ with $|\tilde{x} - x_0| < \delta$ there is a $t_1 > 0$ such that $|\phi_{t_1}(\tilde{x})| = \varepsilon$. Then since m_ε is the minimum of $V(x)$ on S_ε, this would imply that $V(\phi_{t_1}(\tilde{x})) \geq m_\varepsilon$ which contradicts (1.19). Thus for $\tilde{x}$ with $|\tilde{x} - x_0| < \delta$ and $t \geq 0$ it follows that $|\phi_t(\tilde{x})| < \varepsilon$.[34] Note that this is only true if the simplex is invariant under the dynamics of the differential equation system $\dot{x} = F(x)$. The reason is that simplex invariancy implies that the trajectories through $\tilde{x}$ can only leave $\overline{N_\varepsilon(x_0)}$ by crossing S_ε.

So for $\tilde{x}$ with $|\tilde{x} - x_0| < \delta$ and $t \geq 0$, $\phi_t(\tilde{x}) \subset \overline{N_\varepsilon(x_0)}$. Let $\{t_k\}$ be any sequence with $t \to \infty$. Then since $\overline{N_\varepsilon(x_0)}$ is compact, there is a subsequence

[33] We follow the proof of Theorem 3 in Perko (2000, p. 131) and make adjustments to our case where necessary.

[34] By that, we have shown stability of x_0, which is weaker than asymptotic stability.

$\{\phi_{t_n}(\tilde{x})\}$ of $\{\phi_{t_k}(\tilde{x})\}$ that converges to a point $y_0 \in \overline{N_\varepsilon(x_0)}$.[35] Because $V(x)$ is a continuous function, $V(\phi_{t_n}(\tilde{x})) \to V(y_0)$. Since $V(x)$ is strictly decreasing along the trajectories of $\dot{x} = F(x)$ we have that

$$V(\phi_t(\tilde{x})) > V(y_0)$$

for $t \geq 0$. Now we have to determine y_0. Assume that $y_0 \neq x_0$. Then for $s > 0$ we have $V(\phi_s(y_0)) < V(y_0)$. Continuity of $V(x)$ implies that for all y sufficiently close to y_0 we have $V(\phi_s(y)) < V(y_0)$ for $s > 0$. But then for $y = \phi_{t_n}(\tilde{x})$ and n sufficiently large, we have $V(\phi_{s+t_n}(\tilde{x})) < V(y_0)$ which contradicts the above inequality. So by contradiction we have

$$y_0 = x_0\,.$$

Since $V(x)$ is strictly decreasing along trajectories and since the subsequence $\phi_{t_n}(\tilde{x})$ converges to x_0, it follows for every sequence $t_k \to \infty$ that $\phi_{t_k}(\tilde{x}) \to x_0$. Therefore $\phi_t(\tilde{x}) \to x_0$ as $t \to \infty$, which means that x_0 is asymptotically stable. □

We now have to show that there exists a Liapunov function for system (1.7) as defined in Theorem 1. We will give evidence of existence by presenting an example: We show in the following that function $V(x)$,

$$V(x) \;=\; x_1^2 + (1 - x_2)^2\,,$$

is a Liapunov function. It is clear that $V(x) > 0\ \forall x \in \mathbb{R}^2 \setminus (0,1)$, hence $V(x) > 0\ \forall x \in \Sigma_2$. Further $V(0,1) = 0$. Let us now look at $\dot{V}(x)$.

$$\begin{aligned}\dot{V}(x) &= 2x_1\dot{x}_1 + 2(1 - x_2)(-\dot{x}_2)\\ &= 2x_1^2(1 - x_1 - x_2)[(1 - \tau)(w(x) - 1) - x_2(\tau - c(x))]\\ &\quad - 2(1 - x_2)x_2(1 - x_1 - x_2)[(1 - \tau)w(x) + (1 - x_2)(\tau - c(x))]\,.\end{aligned}$$

We analyze $\dot{V}(x)$ in E, the environment of $(0,1)$, which requires to evaluate $w(x)$ and $c(x)$ in E.

$$\begin{aligned}\lim_{x_2 \to 1} w(x) &= \lim_{x_2 \to 1} \frac{\tau}{1 - \tau} \frac{(1 - x_1 - x_2)(1 - x_2)}{x_1 + x_2} \;=\; 0\\ \lim_{x_2 \to 1} c(x) &= \lim_{x_2 \to 1} p(x)\Big((1 - \tau)w(x) + 1\Big) \;=\; p(0,1)\,.\end{aligned}$$

[35] By Bolzano-Weierstrass: Every sequence in a compact set of $\mathbb{R}^n$ has at least one convergent subsequence (e.g. Koenigsberger, 2001, p. 51).

We conclude that for a sufficiently small environment of $(0,1)$ and for $p(0,1) < \tau$ we have

$$\begin{aligned}
\dot V(x) &= \underbrace{2x_1^2(1-x_1-x_2)}_{>0}\,[\underbrace{(1-\tau)(w(x)-1)}_{<0} - \underbrace{x_2(\tau - c(x))}_{>0}] \\
&\quad - \underbrace{2(1-x_2)x_2(1-x_1-x_2)}_{>0}\,[\underbrace{(1-\tau)w(x)}_{>0} + \underbrace{(1-x_2)(\tau-c(x))}_{>0}] \\
&\Rightarrow \quad V(x) < 0\,.
\end{aligned}$$

So if $p(0,1) < \tau$, $(0,1)$ is asymptotically stable.

Critical point $(\mathbf{0}, \bar{\mathbf{x}}_\mathbf{2})$: The critical point $(0,\bar x_2)$ does not necessarily exist for all τ and $p(x_1,x_2)$. Because of $\frac{\partial F_1}{\partial x_2}(0,\bar x_2) = 0$ the eigenvalues of the jacobian $DF(0,\bar x_2)$ are $\frac{\partial F_1}{\partial x_1}(0,\bar x_2)$ and $\frac{\partial F_2}{\partial x_2}(0,\bar x_2)$. By rearranging (1.13) to

$$(1-\tau)w(0,\bar x_2) - x_2(\tau - c(0,\bar x_2)) = c(0,\bar x_2) - \tau\,,$$

we can write

$$\begin{aligned}
\frac{\partial F_1}{\partial x_1}(0,\bar x_2) &= (1-\bar x_2)\Big((1-\tau)w(0,\bar x_2) - \bar x_2(\tau - c(0,\bar x_2))\Big) \\
&= (1-\bar x_2)\Big(c(0,\bar x_2) - 1\Big)\,.
\end{aligned}$$

For the second eigenvalue we note that

$$(1-\tau)\frac{\partial w}{\partial \bar x_2}(0,\bar x_2) \;=\; \tau\frac{\bar x_2^2 - 1}{\bar x_2^2}$$

and can then derive

$$\begin{aligned}
\frac{\partial F_2}{\partial x_2}(0,\bar x_2) &= \bar x_2(1-\bar x_2)\,\cdot \\
&\quad \left[(1-\tau)\frac{\partial w}{\partial x_2}(0,\bar x_2) - (\tau - c(0,\bar x_2)) - (1-\bar x_2)\frac{\partial c}{\partial x_2}(0,\bar x_2)\right] \\
&= \bar x_2(1-\bar x_2)\,\cdot \\
&\quad \left[\tau\left(\frac{\bar x_2^2-1}{\bar x_2^2} - 1\right) + c(0,\bar x_2) - (1-\bar x_2)\frac{\partial c}{\partial x_2}(0,\bar x_2)\right] \\
&= \bar x_2(1-\bar x_2)\left[-\frac{\tau}{\bar x_2^2} + \frac{\tau}{\bar x_2} - (1-\bar x_2)\frac{\partial c}{\partial x_2}(0,\bar x_2)\right] \\
&= \bar x_2(1-\bar x_2)^2\left[-\frac{\tau}{\bar x_2^2} - \frac{\partial c}{\partial x_2}(0,\bar x_2)\right]\,.
\end{aligned}$$

We now have to determine the algebraic sign of the two eigenvalues. The second eigenvalue is positive if

$$\left[-\frac{\tau}{\bar x_2^2} - \frac{\partial c}{\partial x_2}(0,\bar x_2)\right] > 0\,.$$

By replacing $\frac{\partial c}{\partial x_2}(0,\bar x_2)$ by

$$\begin{aligned}
&\frac{\partial p}{\partial x_2}(0,\bar x_2)\,[(1-\tau)w(0,\bar x_2) + 1] + p(0,\bar x_2)(1-\tau)\frac{\partial w}{\partial \bar x_2}(0,\bar x_2) \\
&= \frac{\partial p}{\partial x_2}(0,\bar x_2)\left[\tau\frac{(1-\bar x_2)^2}{\bar x_2} + 1\right] - p(0,\bar x_2)\tau\frac{1-\bar x_2^2}{\bar x_2^2}
\end{aligned}$$

and rearranging terms we find the condition

$$p(0,\bar x_2)\tau(1-\bar x_2^2) - \frac{\partial p}{\partial x_2}(0,\bar x_2)\left[\tau\bar x_2(1-\bar x_2)^2 + \bar x_2\right] \;>\; \tau \qquad (1.20)$$

for a positive second eigenvalue of $DF(0,\bar x_2)$. We have assumed that $p(x_1,x_2)$ is decreasing in x_2, so it is possible to find a $p(x_1,x_2)$ that is consistent with (1.20). The critical point $(0,\bar x_2)$ is then either a saddle (if $\tau < \bar x_2$) or a source (if $\bar x_2 < \tau$). We discuss respective properties of $p(x_1,x_2)$ in the text.

Critical point $(\mathbf{1},\mathbf{0})$: This critical point always exists but it is not hyperbolic. We will show instability by analyzing the system (1.7) in the vicinity of $(1,0)$. This can be done by looking at

$$-\frac{\partial F_1}{\partial x_1}(1,0)\,.$$

The intuition is as follows. Since $(1,0)$ is a fixed point, we have that $\dot x_1(1,0) = F_1(1,0) = 0$. We now check which sign $F_1(x)$ takes if we marginally deviate from $(0,1)$ by decreasing x_1 marginally (and increasing x_2 and x_3 marginally). If $-\frac{\partial F_1}{\partial x_1}(1,0)$ is negative (positive), we know that $\dot x_1$ is negative (positive) in the vicinity of $(1,0)$ since it is zero in $(1,0)$. We find

$$-\frac{\partial F_1}{\partial x_1}(1,0) \;=\; -(-1)\,[(1-\tau)(w(1,0)-1)] \;=\; -(1-\tau)\,.$$

So marginally deviating from $(1,0)$ by marginally decreasing x_1 causes the function F_1 to take a negative value. That means that a solution curve $x(t)$

of system (1.7) starting in the vicinity of $(1,0)$ will move away from $(1,0)$. So $(1,0)$ cannot be asymptotically stable.

Critical point $\{(\mathbf{x_1}, \mathbf{1}-\mathbf{x_1}) \mid \mathbf{x_1} \in]\mathbf{0},\mathbf{1}[\}$**:** This set of critical points always exists and its elements are neither hyperbolic nor isolated critical points. In order to show instability we can use the arguments made for the critical point $(1,0)$. We will check what signs $F_1(x_1,x_2)$ and $F_2(x_1,x_2)$ take in the vicinity of $(x_1, 1-x1)$. Deviating from the edge of the simplex $(x_1, 1-x_1)$ means that we marginally decrease x_1 and x_2 at the same time. So we are interested in the signs of

$$\begin{aligned} &-\frac{\partial F_1}{\partial x_1}(x_1,1-x_1) - \frac{\partial F_1}{\partial x_2}(x_1,1-x_1) \quad \text{and} \\ &-\frac{\partial F_2}{\partial x_1}(x_1,1-x_1) - \frac{\partial F_2}{\partial x_2}(x_1,1-x_1)\,. \end{aligned}$$

Note first that

$$w(x_1,1-x_1) = 0 \quad \text{and} \quad c(x_1,1-x_1) = p(x_1,1-x_1)\,.$$

We use the expressions of we have given for the elements of the Jacobian $DF(x)$ and find

$$\begin{aligned} &- \frac{\partial F_1}{\partial x_1}(x_1,1-x_1) - \frac{\partial F_1}{\partial x_2}(x_1,1-x_1) \\ =\ & -\Big\{ - x_1\Big[(1-\tau)(w(x_1,1-x_1)-1) - (1-x_1)(\tau - c(x_1,1-x_1))\Big]\Big\} \\ & -\Big\{ - x_1\Big[(1-\tau)(w(x_1,1-x_1)-1) - (1-x_1)(\tau - c(x_1,1-x_1))\Big]\Big\} \\ =\ & 2x_1\Big[-(1-\tau) - (1-x_1)(\tau - p(x_1,1-x_1))\Big] \\ =\ & -2x_1\Big[1 - p(x_1,1-x_1) - x_1(\tau - p(x_1,1-x_1))\Big]\,. \end{aligned}$$

If $\tau > p(x_1,1-x_1)$ then $1-p(x_1,1-x_1) > \tau - p(x_1,1-x_1) > 0$ and hence $1-p(x_1,1-x_1) - x_1(\tau - p(x_1,1-x_1)) > 0$. If $p(x_1,1-x_1) > \tau$ then $0 > x_1(\tau - p(x_1,1-x_1))$ and again $1-p(x_1,1-x_1)-x_1(\tau-p(x_1,1-x_1)) > 0$. If $p(x_1,1-x_1) = \tau$ then $1-p(x_1,1-x_1) - x_1(\tau - p(x_1,1-x_1)) = 0$ again. So we can state

$$-\frac{\partial F_1}{\partial x_1}(x_1,1-x_1) - \frac{\partial F_1}{\partial x_2}(x_1,1-x_1) \ < \ 0\,.$$

That means that all solution curves starting in the vicinity of $(x_1, 1-x_1)$ will move away from $x_1 = 1$. Because this result holds for all $x_1 \in (0,1)$ the critical points in the set $\{(x_1,1-x_1) \mid x_1 \in]0,1[\}$ cannot be stable; it is impossible that trajectories starting in the vicinity of this edge of the simplex move towards a point in this set. It is redundant to determine the sign of $F_2(x)$ in the vicinity of $\{(x_1,1-x_1) \mid x_1 \in]0,1[\}$.

Critical point $(\tau, \mathbf{0})$**:** This critical point does always exist. It is easy to see that $\frac{\partial F_2}{\partial x_1}(\tau,0) = 0$ since $x_2 = 0$. So we know that the two eigenvalues of $DF(\tau,0)$ are $\frac{\partial F_1}{\partial x_1}(\tau,0)$ and $\frac{\partial F_2}{\partial x_2}(\tau,0)$. Note that

$$\begin{aligned} w(\tau,0) &= 1\,, \\ \frac{\partial w}{\partial x_1}(\tau,0) &= -\frac{1}{\tau(1-\tau)}\,, \\ c(\tau,0) &= (2-\tau)p(\tau,0)\,. \end{aligned}$$

The eigenvalues of $DF(\tau,0)$ are

$$\begin{aligned} \frac{\partial F_1}{\partial x_1}(\tau,0) &= -(1-\tau) \\ \frac{\partial F_1}{\partial x_1}(\tau,0) &= (1-\tau)[1-(2-\tau)p(\tau,0)]\,. \end{aligned}$$

We conclude that if $\frac{1}{2-\tau} < p(\tau,0)$, the critical point is a sink. If $p(\tau,0) < \frac{1}{2-\tau}$, the critical point is a saddle.

Critical point $(\tau - \hat{\mathbf{x}}_\mathbf{2}, \hat{\mathbf{x}}_\mathbf{2})$**:** In order to analyze the stability of $(\tau - \hat{x}_2, \hat{x}_2)$, we evaluate the elements of $DF(\tau-\hat{x}_2,\hat{x}_2)$. Note that

$$\begin{aligned} w(\tau-\hat{x}_2,\hat{x}_2) &= 1-\hat{x}_2 \\ \frac{\partial w}{\partial x_1}(\tau-\hat{x}_2,\hat{x}_2) &= -\frac{1-\hat{x}_2}{\tau(1-\tau)} \\ \frac{\partial w}{\partial x_2}(\tau-\hat{x}_2,\hat{x}_2) &= -1-\frac{1-\hat{x}_2}{\tau(1-\tau)}\,. \end{aligned}$$

We find

$$\begin{aligned} \frac{\partial F_1}{\partial x_1}(\tau-\hat{x}_2,\hat{x}_2) &= (1-\tau)(\tau-\hat{x}_2)\Big(-\frac{1-\hat{x}_2}{\tau} + \hat{x}_2\frac{\partial c}{\partial x_1}(\tau-\hat{x}_2,\hat{x}_2)\Big) \\ \frac{\partial F_1}{\partial x_2}(\tau-\hat{x}_2,\hat{x}_2) &= (1-\tau)(\tau-\hat{x}_2)\Big(-\frac{1-\hat{x}_2}{\tau} + \hat{x}_2\frac{\partial c}{\partial x_2}(\tau-\hat{x}_2,\hat{x}_2)\Big) \end{aligned}$$

$$\frac{\partial F_2}{\partial x_1}(\tau - \hat{x}_2, \hat{x}_2) = (1-\tau)\hat{x}_2\Big(-\frac{1-\hat{x}_2}{\tau} - (1-\hat{x}_2)\frac{\partial c}{\partial x_1}(\tau - \hat{x}_2, \hat{x}_2)\Big)$$
$$\frac{\partial F_2}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2) = (1-\tau)\hat{x}_2\Big(-\frac{1-\hat{x}_2}{\tau} - (1-\hat{x}_2)\frac{\partial c}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2)\Big).$$

Because we have assumed that $\frac{\partial p}{\partial x_2} < 0$ we have $\frac{\partial c}{\partial x_2} < 0$, and therefore we know that

$$\frac{\partial F_1}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2) < 0\,.$$

Again from Assumption 1 we derive

$$\frac{\partial F_1}{\partial x_1}(\tau - \hat{x}_2, \hat{x}_2) > \frac{\partial F_1}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2)$$
$$\frac{\partial F_2}{\partial x_1}(\tau - \hat{x}_2, \hat{x}_2) > \frac{\partial F_2}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2)\,.$$

Let us abbreviate the elements of $DF(\tau - \hat{x}_2, \hat{x}_2)$ by j_i, such that

$$DF(\tau - \hat{x}_2, \hat{x}_2) = \begin{pmatrix} j_1 & j_2 \\ j_3 & j_4 \end{pmatrix}.$$

The standard formula for eigenvalues of 2×2-matrices yields for $DF(\tau - \hat{x}_2, \hat{x}_2)$

$$\lambda_1 = \frac{1}{2}\Big(j_1 + j_4 + \sqrt{j_1^2 - 2j_1j_4 + j_4^2 + 4j_2j_3}\Big),$$
$$\lambda_2 = \frac{1}{2}\Big(j_1 + j_4 - \sqrt{j_1^2 - 2j_1j_4 + j_4^2 + 4j_2j_3}\Big).$$

We are only interested in the real parts of the eigenvalues, for they determine the stability of the critical point. A square root of a real discriminant Δ always either has a zero real part (if $\Delta \leq 0$) or a positive real part (if $\Delta > 0$). Therefore we have that $\lambda_1 \geq \lambda_2$. If $(\tau - \hat{x}_2, \hat{x}_2)$ is stable if and only if $\lambda_1 < 0 \wedge \lambda_2 < 0$. So the necessary condition for stability of $(\tau - \hat{x}_2, \hat{x}_2)$ is

$$\begin{aligned} \lambda_1 &< 0 \\ \frac{1}{2}\Big(j_1 + j_4 + \sqrt{j_1^2 - 2j_1j_4 + j_4^2 + 4j_2j_3}\Big) &< 0 \\ j_1 + j_4 &< -\sqrt{j_1^2 - 2j_1j_4 + j_4^2 + 4j_2j_3} \\ (j_1+j_4)^2 &> j_1^2 - 2j_1j_4 + j_4^2 + 4j_2j_3 \\ j_1j_4 &> j_2j_3\,. \end{aligned}$$

We substitute the explicit elements of $DF(\tau - \hat{x}_2, \hat{x}_2)$ into the last equation. Simplifying the expression leaves us with

$$\frac{1}{\tau}(1-\tau)^2(\tau - \hat{x}_2)\hat{x}_2(1-\hat{x}_2)\left(\frac{\partial c}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2) - \frac{\partial c}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2)\right) > 0$$
$$\Rightarrow \quad \frac{\partial c}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2) > \frac{\partial c}{\partial x_2}(\tau - \hat{x}_2, \hat{x}_2)\,.$$

This is contradictory to Assumption 1. Thus, as long as we adhere to Assumption 1, $(\tau - \hat{x}_2, \hat{x}_2)$ cannot be stable.

Table 1.2: Critical points of the corruption game.

Critical point (x_1, x_2)	Conditions on τ and $p(x_1, x_2)$ for existence of critical point	Asymptotic stability
$(0,0)$	none	unstable
$(0,1)$	none	asymptotically stable if $p(0,1) < \tau$
$(0,\bar{x}_2)$	$\bar{x}_2$ such that $p(0,\bar{x}_2) = \frac{\tau}{\tau(1-\bar{x}_2)^2+\bar{x}_2}$	asymptotically stable if $\tau < -\bar{x}_2^2\frac{\partial c}{\partial x_2}(0,\bar{x}_2)$
$(1,0)$	none	unstable
$(x_1, 1-x_1)$ with $x_1 \in]0,1[$	none	unstable
$(\tau, 0)$	none	asymptotically stable if $\frac{1}{2-\tau} < p(\tau,0)$
$(\bar{\tau} - \hat{x}_2, \hat{x}_2)$	$\hat{x}_2$ such that $p(\tau - \hat{x}_2, \hat{x}_2) = \frac{1}{(1-\tau)(1-\hat{x}_2)+1}$	unstable

Global Theory of Nonlinear Systems

It is left to show that no other attracting sets exist [36] and attractors [37] than those found in the last subsection.

The Generalized Poincaré-Bendixson Theorem for Analytic Systems states that the ω-limit set [38] of any trajectory of a two-dimensional, relatively-prime, analytic system is either a critical point, a cycle, or a compound separatrix cycle. We show below that the solution trajectories of system (1.7) cannot be closed. This allows us to exclude limit cycles and compound separatrix cycles as ω-limits. We will then be able to conclude that only evolutionary equilibria can be attractors.

In order to show that our system does not have any closed trajectories, we apply index theory, a method that describes global behavior of the solutions to a differential equation system (see e.g. Strogatz, 1994, Chapter 6). In Proof 1.A.5 we have calculated all critical points of system (1.7). They all are on the boundary of the simplex except $(\tau - \hat{x}_2, \hat{x}_2)$, which is a saddle. We now assume that there exists a closed trajectory to (1.7). Figure 1.5 shows all qualitatively different locations a closed trajectory could occupy, they are indicated by the dotted curves T_1, T_2, and T_3. The index at each of the critical points is also shown in the figure (for an explanation of how to calculate the index at critical points, see Strogatz, 1994, Chapter 6).
We can rule out closed trajectories as follows. Trajectories like T_1 are impossible because they cross the boundary of the simplex. The reason is the following. From Proposition 3 we know that the simplex boundary is invariant under system (1.7). So the boundary of the simplex contains straight-line trajectories. Since trajectories cannot cross,[39] we can exclude trajectories like T_1. Trajectories like T_2 can be excluded as well because they do not enclose any fixed points at all. And trajectories like T_3 violate

[36] A closed invariant set $A \in E$ is called an *attracting set* of a system $\dot{x} = F(x)$ if there is some neighborhood U of A such that for all $x \in U$, $\phi_t(x) \in U$ for all $t \geq 0$ and $\phi_t(x) \to A$ as $t \to \infty$.

[37] An *attractor* is an attracting set containing a dense orbit (a dense orbit is an orbit that comes arbitrarily close to each point in the attractor).

[38] A point $p \in E$ is an ω-limit point of the trajectory $\phi(t, x_0)$ of the system $\dot{x} = F(x)$ if there is a sequence $t_n \to \infty$ such that $\lim_{n\to\infty} \phi(t_n, x) = p$.

[39] This follows directly from the Fundamental Existence-Uniqueness Theorem.

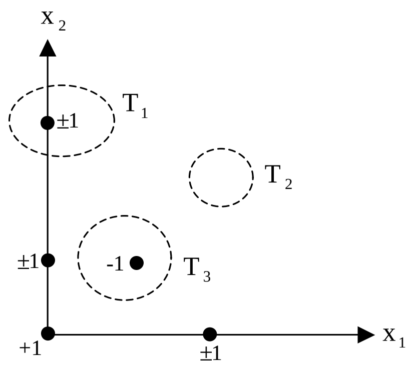


Figure 1.5: Locations of closed trajectories.

the requirement that the indices inside the closed trajectories must sum up to 1 (see e.g. Strogatz, 1994, Theorem 6.8.2., p. 180). We conclude that system (1.7) does not have any closed trajectories. Consequently, the ω-limit set of any trajectory of our system is a critical point. It is clear from the definition of an attractor and the concept of asymptotic stability, that only an asymptotically stable critical point can be an attractor. We conclude that almost every trajectory through a point $x \in \Sigma$ approaches an EE in the limit. The sole exceptions are those trajectories that are the separatrices of the system.

From the Poincaré-Bendixson Theorem we know that if a trajectory of a planar system is confined to a closed, bounded region, then the trajectory is either attracted by a critical point or a closed trajectory. Since the simplex is invariant under the dynamics of system (1.7) and since we have shown that system (1.7) has no closed trajectories, we can conclude, that there always exists an attractor in the corruption game. This results holds for all $p(x)$ and all τ.

1.B Appendix: Games with two Strategies

In this Appendix, we analyze frequency dependent evolutionary games with two strategies. From Proposition 4 we know that we can write a the payoff matrix of a two-strategy frequency dependent evolutionary game by the matrix

$$B(x) = \begin{pmatrix} a(x) & 0 \\ 0 & b(x) \end{pmatrix},$$

where $a(x)$ and $b(x)$ are Lipschitz-continuous by assumption. The replicator dynamics (1.4) can be written as

$$\begin{aligned} \dot{x}_1 &= a(x)x_1 - \big(a(x)x_1^2 + b(x)x_2^2\big) \\ \dot{x}_2 &= b(x)x_2 - \big(a(x)x_1^2 + b(x)x_2^2\big) . \end{aligned}$$

By Proposition 3, $\dot{x}_2 = -\dot{x}_1$ and we therefore concentrate on only one equation. We substitute x_2 by $1 - x_1$ which leaves us with

$$\begin{aligned} \dot{x}_1 &= a(x)x_1 - \big(a(x)x_1^2 + b(x)(1-x_1)^2\big) \\ &= x_1(1-x_1)\,(x_1 a(x_1) - (1-x_1)\,b(x_1)) \;=\; F(x_1) . \end{aligned}$$

For arbitrary functions $a(x)$ and $b(x)$ the replicator dynamics can yield very complicated behavior because $F(x_1)$ may have many critical points which qualify for EE. It is therefore interesting to state a theorem concerning the number of critical points. Proposition 6 assumes that $a(x)$ and $b(x)$ are polynomials. Polynomials are a proper subset of the set of Lipschitz-continuous functions, so our proposition does not apply for all games.[40]

Proposition 6 *Assume that $a(x)$ and $b(x)$ are polynomials while one of them has a non-zero degree. The number of interior EE in a two-strategy frequency dependent evolutionary game with replicator dynamics is equal to or smaller than the degree of the polynomial with the higher degree.*

[40] However, the Weierstrass polynomial approximation theorem assures us that there is a sequence of polynomials which converges uniformly to the function $F(x_1)$ on $[0,1]$. This means that we can find a polynomial that approximates $F(x_1)$ on $[0,1]$ to any desired degree of accuracy.

We prove this Proposition at the end of this Appendix.

Apart from the interior EE, the games can additionally have one or two EE at $x_1 = 0$ (if $F(0) \leq 0$) and $x_1 = 1$ (if $F(1) \geq 0$).

Most economic interpretations will not need highly nonlinear payoff functions. For many applications it will be sufficient to model whether a payoff increases or decreases with the frequency of a strategy, and if these changes become stronger or weaker the higher the frequency of the strategy. Even with quite simple payoffs, the number of critical points of $F(x_1)$ cannot be determined generally. The only class of payoff functions that allows for a more precise description of equilibrium behavior is the class of linear functions. We can derive the result from Proposition 6 because linear functions are polynomials of degree one.

Lemma 1 *If $a(x)$ and $b(x)$ are linear functions, a two-strategy frequency dependent evolutionary game with replicator dynamics has at most one interior EE and at most one unstable critical point.*

We now compare the standard games with their frequency dependent counterparts. Of the former, we distinguish between three categories: Prisoners' Dilemma (Type I and II), Coordination Games, and Hawk-Dove Games (see Weibull, 1995, p. 75). Analogous to these categories we assume for the frequency dependent Prisoners' Dilemma I $a(x) > 0$ and $b(x) < 0$, for the frequency dependent Prisoners' Dilemma II $a(x) < 0$ and $b(x) > 0$, for the frequency dependent Coordination Game $a(x) > 0$ and $b(x) > 0$, and for the frequency dependent Hawk-Dove Game $a(x) < 0$ and $b(x) < 0$. Table 1.3 summarizes our findings, which are proved in the end of this Appendix. We see that no matter how the payoffs in a Prisoners' Dilemma change with the frequency of a strategy, the EE will be the same as in a game with constant payoffs. However, for Coordination Games the situation is different: depending on the payoff function, it is now possible that $x_1 = 1$ is no longer an EE, or that it is replaced by an EE in the interior of the strategy space. In the case of the Hawk-Dove Game, frequency dependent payoffs can change the dynamics of the game too. While the game with constant payoffs featured a unique interior EE, the frequency dependent game can either have a (different) unique interior EE, too or have an EE at $x_1 = 1$, or both.

Table 1.3: Evolutionary Equilibria of standard and FD-games.

Game Category	EE of Standard Game	EE of FD-Game
PD I	$x_1 = \{0\}$	$x_1 = \{0\}$
PD II	$x_1 = \{1\}$	$x_1 = \{1\}$
CG	$x_1 = \{0, 1\}$	$x_1 = \{0\}$ or $x_1 = \{0, 1\}$ or $x_1 = \{0, p_{cg}\}$
HD	$\{x_1 = \frac{b}{a+b}\}$	$x_1 = \{1\}$ or $x_1 = \{p_{hd}, 1\}$ or $x_1 = \{p_{hd}\}$

Our extension to FD-games allows to conjoin the different categories of games. We demonstrate this with the next example.

Example 1
Agents parking their cars have the choice to pay the official fee of 1 (Strategy 1) or to park illegally (Strategy 2). The expected fine for illegal parking decreases with the number of illegally parked cars (e.g. because officers do not manage to make out tickets for all illegally parked cars in a given time interval). We assume that the expected fine is $2x_1(t)$. Then the payoff matrix is

$$A(x) = \begin{pmatrix} -1 & -1 \\ -2x_1(t) & -2x_1(t) \end{pmatrix}.$$

For $x_1 = 0$ this game is a Prisoners' Dilemma I, for $x_1 = 1$ it is a Prisoners' Dilemma II. The set of EE is $\{x_1 = 0, x_1 = 1\}$. If we start off with less (more) than half of the agents paying the fee, nobody (everybody) pays in the equilibrium. While if we would have a fixed expected fine we would observe equilibria that are independent of the initial condition: either $x_1 = 0$ for an expected fine smaller than one, or $x_1 = 1$ for an expected fine greater than one.

Proof of Proposition 6

We define the function $f(x_1)$ as

$$\begin{aligned} f(x_1) &= x_1 a(x_1) - (1 - x_1)\, b(x_1) \quad \text{which implies} \\ F_1(x) &= x_1(1 - x_1) f(x_1)\,. \end{aligned}$$

The degree formulas for the polynomial ring imply

$$\begin{aligned} \deg\,(xa(x)) &= \deg a(x) + 1 \quad \text{and} \\ \deg\,((1-x)b(x)) &= \deg b(x) + 1 \\ \Rightarrow \quad \deg f(x) &= \max\{\deg a(x), \deg b(x)\} + 1\,. \end{aligned}$$

By the fundamental theorem of algebra, the polynomial $f(x)$ has $\deg f(x)$ roots. So it has at most $\deg f(x)$ roots on the interval $[0, 1]$. We have assumed that either $\deg a(x) > 0$ or $\deg b(x) > 0$, so $\deg f(x) > 1$. Therefore, $f(x)$ has potentially two or more roots. Not all of these potential roots can be EE, because we need $f'(x) < 0$ for a root to be an EE. The number of potential roots will be reduced by 1 at least. Thus we can say that $f(x)$ has at most as many roots as the polynomial with the higher degree has. Since $x_1(1 - x_1) > 0$ on $]0, 1[$, the same is true for $F(x_1)$.

Calculations for Table 1.3

Prisoners' Dilemma I: From $a(x_1) < 0$ and $b(x_1) > 0$ we have that $f(x_1) = xa(x_1) - (1 - x_1)b(x_1) < 0$ and the system has thus no critical point and $\dot{x}_1$ is negative on $[0, 1]$. Thus, $x = 0$ is an EE.

Prisoners' Dilemma II: From $a(x_1) > 0$ and $b(x_1) < 0$ we have that $f(x_1) = xa(x_1) - (1 - x_1)b(x_1) > 0$ and the system has thus no critical point and $\dot{x}_1$ is positive on $[0, 1]$. Thus, $x = 1$ is an EE.

Coordination Game: The Coordination Game requires $a(x_1) = a_1 + a_2 x > 0$ and $b(x_1) = b_1 + b_2 x > 0$. Hence, $a_1 > 0$ and $b_1 > 0$. For $F(x_1)$ we find

$$\begin{aligned} F(x_1) &= a_1 x_1 + a_2 x_1^2 - (1 - x_1)(b_1 + b_2 x_1) \\ &= (a_2 + b_2)x_1^2 + (a_1 + b_1 - b_2)x - b_1 \end{aligned}$$

The roots of $F(x_1)$ are

$$r_{1,2} = \frac{-(a_1 + b_1 - b_2) \pm \sqrt{(a_1 + b_1 - b_2)^2 + 4(a_2 + b_2)b_1}}{2(a_2 + b_2)}\,.$$

Note that $F(0) = -b_1 < 0$. So if $r_1 \in]0, 1[$ and $r_2 \in]0, 1[$, then $x_1 = 0$ and $x_1 = r_2$ are EE. We have named r_2 as p_{cd} in Table 1.3. If either $r_1 \in]0, 1[$

or $r_2 \in]0,1[$ but not both, then this root is an unstable critical point and $\{x_1 = 0, x_1 = 1\}$ are the EE. If none of the roots lies in $]0,1[$, then $F(x_1) < 0$ on $]0,1[$ and $x_1 = 0$ is the only EE.

Hawk-Dove Game: We proceed analogously to the calculation for the coordination game. In a Hawk-Dove Game, the FD-payoffs satisfy $a(x_1) = a_1 + a_2 x < 0$ and $b(x_1) = b_1 + b_2 x < 0$. It follows that $a_1 < 0$ and $b_1 < 0$. Note that $F(0) = -b_1 > 0$ (check function above). The roots are the same as well. If $r_1 \in]0,1[$ and $r_2 \in]0,1[$, then $x_1 = r_1$ and $x_1 = 1$ are EE. We have named r_1 as $p_{hd,1}$ in Table 1.3. If either $r_1 \in]0,1[$ or $r_2 \in]0,1[$ but not both, then this root is the sole EE. We have named that root $p_{hd,2}$ in Table 1.3. If none of the roots lies in $(0,1)$, then $F(x_1) > 0$ on $(0,1)$ and $x_1 = 1$ is the only EE.

Chapter 2

Imitating Illegal Activities - A Spatial Model With Heterogeneity

> The problem of a rational economic order is determined precisely by the fact that the knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess...
>
> Friedrich A. von Hayek

2.1 Introduction

In Chapter 1 we modelled the dynamics of corruption as a frequency-dependent evolutionary game with replicator dynamics. We based the latter on a standard imitation rule as well as on the assumption that payoffs of strategies could only be observed with some noise. However, the introduction of noisy observations served as a shortcut for comprehending the agents' difficulties in obtaining information. In the present chapter, we explicitly focus on the informational aspect of strategic interaction of legal and illegal strategies. Instead of applying a shortcut as in Chapter 1, we now model the flow of information explicitly and consider the special nature of this process. In order to concentrate on the impact that the flow of information has on game dynamics, we abandon the complex structure of the corruption game. In the following we present a simple evolutionary game with two strategies only, the legal one and the illegal one. As in the previous chapter, we assume that players make their strategy decision by imitating more successful strategies. The imitation rule is discussed below. Note that we aim to describe the spread of illegal activities in a population. This is done by analyzing the absorbing states of the imitation dynamics if just one agent acts illegally in the initial period.

Before we proceed to the detailed analysis, let us first provide a motivation for our model by illustrating that imitation is a special form of social interaction. In the next section, we firstly define the term *social interaction* generally, then point out why social interactions are relevant for the study of illegal activities, and briefly explain the kind of social interactions to which imitation belongs. Secondly, we discuss the differences between information on legal activities and information on illegal activities. We identify two basic differences and then go on to show how we can include them in our model. Finally, we present our model, the results and the related literature.

2.2 Preliminary Considerations

2.2.1 Imitation as Social Interaction

Imitation as social interaction. Gary Becker was the first to apply economic theory to social issues such as crime. His rational choice approach later became the standard framework for economists in studying illegal activities. Thereby the engagement in illegal activities is understood as follows (see Becker, 1968): An individual decides to act illegally if the expected cost of the illegal activity is lower than the expected benefits. This implies that exogenously increased costs of crime will imply lower crime rates.

Although this result is consistent with the findings of empirical studies (e.g. Ehrlich, 1973), the high variance of crime rates across time and space cannot be explained with this approach and remains one of the oldest puzzles in social sciences (see Glaeser et al., 1996).

In order to understand the circumstances under which illegal activities

can arise and persist, it is important to know all the factors that additionally or alternatively influence the decision of the individual. Many researchers identify social interaction as an important determinant of criminal activity.

Social interaction is generally defined as the impact of one individual's action on the actions of others. Models of social interaction explaining variations in the crime rate have included spatial (e.g. Glaeser et al., 1996) as well as dynamic models (e.g. Sah, 1991). In an experiment testing for criminal behavior, Falk and Fischbacher (2002) find that almost half of the persons under examination behave conditionally on their environment, i.e. they are more likely to commit criminal acts in a high crime environment.

An individual's actions can affect other individuals' actions in different ways. Manski (2000) identifies three channels of social interaction: constraints, expectations and preferences. *Expectations interactions* imply that individuals facing a decision form expectations concerning an action's outcome by observing the actions chosen and outcomes experienced by others. This chapter focuses on this form of social interaction by adopting an imitation rule. However, in contrast to the existing literature, we consider the special nature of information concerning illegal activities when modelling expectations interactions.

2.2.2 Informational Distortions

Information on illegal activities differs in two ways from information on legal activities: The first specific attribute of information on illegal activities is that it must be hidden and is thus hard to obtain. We therefore say that information on illegal activities is *scarce*.

The reason why knowledgeable individuals prefer to hide their information is twofold: They do not want to reveal involvement in illegal activities in the first place nor do they wish to be prosecuted for offering services empowering others to engage in an illegal activity. As a consequence, agents have to rely on secretly passed on knowledge and are only willing to disclose any information if they consider their counterpart to be trustworthy.[1]

This leads us to the second specific attribute of information on illegal

[1] This feature of our model can also be interpreted differently. The disclosure of information need not occur deliberately. We could also envisage situations where agents are incapable of hiding their activities and outcomes from all their opponents.

activities. When deciding whether or not to engage in an illegal activity, individuals have to put up not only with scarce information but also with private information. In the case of an illegal activity, individuals miss out on research results, expert reports, and user evaluations. They must rely on what they learn in secrecy. In the absence of any large sample results, agents will not be able to detect distortions or failings in this information due to personal factors. Therefore we say that information on illegal activities is *non-verifiable*.

In this analysis, we include both informational distortions – scarcity and non-verifiability. Scarcity is included by choosing a spatial model and claiming that only agents located next to each other exchange information. That is, spatial neighbors are assumed to be confidants.[2] In the model, we allow for small and large groups of confidants, which we denote by local and global information. In order to include non-verifiability in our model, we assume different types of agents. An agent's type affects his payoff but cannot be detected by himself or any other agent. The heterogeneity of agents leads to possible over- and underestimating of the illegal activity by the agents.

2.2.3 The Theoretical Approach

We consider an evolutionary game in which a finite number of infinitely lived agents are matched pairwise to play a 2×2 stage game. The stage game has the following structure: The two agents compete for a prize w by choosing either the legal activity (playing fair) or the illegal activity (cheating). Agents are of two types and are nicknamed *high types* and *low types*. High types have a natural advantage over low types: they obtain the prize w with certainty if they meet a low type and if both use the same strategy. If, however, a low type cheats against a fair playing high type, the low type beats the high type with certainty and receives the prize w. If two players of the same type meet and both use the same strategy, they share the prize w. Cheating is costly with cost c satisfying $0 < c < w$.

As our agents are heterogenous, there are two stage games. Firstly, if

[2]Either confidants or observers that cannot be held off, see Footnote 1. Lambsdorff (2002) discusses the reasons for agreements between individuals not to denounce each other in the case of corruption.

two agents of the same type meet, the stage game is a Prisoner's Dilemma, where acting illegally (or playing fair) is the dominant strategy if the cost of cheating is sufficiently low (or high respectively). Secondly, if two agents of opposite type meet, the stage game is an asymmetric game similar to a matching penny game (e.g. Kreps, 1990). In this game, the low type's best response is the strategy that is not being used by the high type, and the high type's best response is the strategy that the low type is using.

In this chapter, we analyze the absorbing states of the imitation dynamics when initially all but one agent are playing fair. Note that the agent who has the illegal strategy at his disposal in the first period, is called the *innovator*. Under what circumstances can an illegal activity spread in a population of agents competing fairly?

There are two reasons why we we limit our considerations to the special class of initial strategy states in which only one agent acts illegally. Firstly, only by restricting the classes of initial strategy states, a complete analysis of the imitation dynamics is practicable for some selected type distributions. Secondly, there are many applications that suggest a single innovator of illegal activities. One application of this setup are illegal technologies that require a profound knowledge to be developed. It is unlikely that a challenging invention is made simultaneously by different agents of a population.[3] Uncertainty can be another reason why there is just one innovator of an illegal technology. If there are high costs involved that are hardly assessable up-front, or if there is uncertainty about the effectuality or the detection probability of the illegal activity, it needs an extraordinarily risk-loving agent that tests the illegal activity. However, it is natural to assume that such agents are rare.

In each period every agent is matched sequentially to all other agents to play the stage game, i.e. agents interact globally. At the end of the period, each agent observes the strategies and average payoffs of a subset of all agents, called the *information set* of an agent. In the following period, agents imitate the strategy with the highest average payoff in their infor-

[3]A demonstrative example of this setup is the spread of a new doping substance in sports, whereby athletes with different talents may learn how to use this substance to improve their winning probabilities. Other applications are sophisticated software system for password theft on the internet, corruption, or loopholes in international financial controls of anti-money laundering agencies.

mation set.[4] To formalize these sets, we locate the agents on a circle as in Ellison (1993).

In most parts of the chapter, we focus on two information settings: local information and global information. With local information, agents observe the strategies and payoffs of their immediate neighbors on the circle only. If agents have global information, they observe the strategies and payoffs of all agents.[5] In other words: the first informational distortion for illegal activities, i.e. scarcity, is captured by local information in our model. Global information depicts a situation in which all agents would share their knowledge about the illegal strategy.

Now let us turn to the second informational distortion, which we referred to as non-verifiability. As discussed above, we capture this aspect by assuming heterogeneity of agents. With heterogenous agents there are four crucial factors that determine the spread of the illegal strategy: the location of the innovator, the type of the innovator, the distribution of types on the circle, and the information available to the agents. These factors determine whether the innovator is able to infect his neighbors and eventually to contaminate the entire population. Since many different distributions of types on the circle are feasible - each of them having potentially different implications for the absorbing states - we focus on two polar cases: maximal segregation on the one hand and minimal segregation on the other hand. In a maximally segregated population, high types and low types are located in two clusters so that there are only two players of each type that have a

[4]In the case of corruption, government employees share information about the expected income from bribes. Note that only corrupt government employees can provide this information because all others do not know about the chances of being caught or the amount of money that can be asked for as a bribe. The imitation rule implies a certain inertia in the behavior of agents, which can be interpreted as cautiousness: As long as an agent does not know the expected payoff of a strategy, he will not choose to adopt the strategy. Kandori et al. (1993) discuss in more detail the assumption that uncertainty leads to inertia. We would like to mention another interpretation of the imitation rule here. If an agent does not adopt a strategy that he does not observe, the reason can also be that he is incapable of doing so. If an illegal activity involves a technology, an agent will need to be in contact with someone who has the technology at his disposal. Doping in sports is an example of an illegal technology.

[5]In sports, for example tennis, all agents compete against all other agents (global interaction). Nevertheless, the players for obvious reasons only share information about the use of illegal performance-enhancing drugs with their best mates (local information).

neighbor of the opposite type. In a minimally segregated population, each agent has two neighbors of opposite type.

How can the results we observe from the model with heterogenous agents and local information, i.e. the model which depicts scarcity and non-verifiability of information for illegal activities, be summarized? In short, there are three important outcomes. Firstly, the population is more likely to end up in an absorbing state whereby all agents cheat if the innovator is a low type. Secondly, if the innovator is a low type, then a minimally segregated population is more resistant to the illegal strategy than a maximally segregated population. The reason for this result is that in a minimally segregated population, each low type (including the innovator) is surrounded by high types, who are less prone to imitate the illegal strategy. Thirdly, by contrast, if the innovator is a high type, a maximally segregated population is more resistant to the illegal strategy than a minimally segregated population. The reason for this result is that the innovator is again surrounded by high types, who are less prone to imitate the illegal strategy.

What is the role of information for these results? In order to answer this question, we describe the effect of local and global information on the absorbing states. First, with global information, in contrast to local information, the location and the type of the innovator of the illegal strategy and the distribution of types on the circle are irrelevant. Second, local information reduces the spread of the illegal strategy if agents are minimally segregated relative to a situation where agents have global information. For a maximally segregated population this result is only true if the innovator is a high type. Third, with local information some agents under- and some overestimate the true benefit of the illegal strategy. There is no such effect with *complete* information where each player type knows the true benefit of each strategy for his type.

Note that our analysis is most closely related to the work of Ellison (1993), Eshel et al. (1998), and Kandori et al. (1993). Kandori et al. (1993) consider the limiting distribution when individual mutation rates go to zero for the class of 2×2 stage games. The players' period payoffs are the expected values of the stage game given the (distribution of) strategy choices of all players. As in Kandori et al. (1993) we assume "global interaction." Ellison (1993) investigates the limiting distributions and the speed of convergence in a similar model as Kandori et al. (1993). The crucial difference between his and the Kandori et al. (1993) approach is, however, that players interact and obtain information locally.

In short, Kandori et al. (1993) investigate global interaction and global information, while Ellison (1993) focusses on local interaction and local information. In contrast to both of them, we combine local information with global interaction. Moreover, we introduce heterogeneity among agents, which generates different stage games. Finally, we adopt the imitation rule of Eshel et al. (1998) where players can only play the strategies they observe in their information set. Like Ellison (1993), Eshel et al. (1998) may be classified as a local interaction and local information game.

The rest of this chapter is organized as follows. In Section 2.3, we describe the basic model with homogenous agents. Sections 2.4 and 2.5 analyze the model with heterogenous agents and global and local information, respectively. In Section 2.6.1, we allow for mutations. Section 2.7 covers some concluding remarks. We direct all proofs of Propositions to the Appendix 2.A.

2.3 Homogeneity of Agents

In this section, we set up the model with homogeneous agents. We first describe the stage game, specify how agents are located and how they adopt or choose new strategies. We then analyze the absorbing strategy states.

2.3.1 The Stage Game

We consider a finite population with $N > 1$ infinitely living agents denoted by $i = 1, ..., N$. In each period t, every agent is sequentially matched to all other agents to play a 2×2 stage game. In each stage game, the agents compete for a prize w. The strategy space is $\{C, D\}$, where C stands for playing "clean", i.e. fair, and D for cheating, i.e. playing "doped." Within a period, a player cannot change his or her pure strategy. Furthermore, mixed strategies are ruled out.

The payoffs of the stage game are as follows. If both agents play strategy C, each receives $\frac{w}{2}$. If both play D, each gets $\frac{w}{2} - c$, where $c \in (0, w)$ is the cost of cheating. This reflects the fact that each agent prefers to obtain the prize w by playing C rather than by using D. Finally, if an agent plays

D against a clean player, he gets $w - c$ and the clean player 0. Thus, the payoff matrix is

$$A = \begin{pmatrix} \frac{w}{2} & 0 \\ w - c & \frac{w}{2} - c \end{pmatrix}. \tag{2.1}$$

The stage game defined in (2.1) is a Prisoner's Dilemma (see Weibull, 1995). If $c \lessgtr \frac{w}{2}$, D, respectively C, is the dominant strategy.

Agent i's period payoff in period t, $u_{i,t}$, is the average payoff from the $N - 1$ matches,

$$u_{i,t}(\sigma_{i,t}, \sigma_{-i,t}) = \frac{1}{N-1} \sum_{-i} a(\sigma_{i,t}, \sigma_{-i,t})$$

where $\sigma_{i,t}$ is his strategy in period t, $\sigma_{-i,t}$ are the strategies chosen by all other players in period t, and the payoffs $a(.,.)$ are the corresponding elements of A in (2.1).

2.3.2 Location and Imitation

In order to model incomplete information, we assume that agents are located on a circle on the positions $1, 2, 3, ..., N$. In each period, each agent i obtains information about the period payoffs and the strategies chosen by the agents $i \pm k$ (modulo N) with $k \in \{1, ..., \lfloor \frac{N}{2} \rfloor\}$, where $\lfloor \frac{N}{2} \rfloor$ is the largest integer $\leq \frac{N}{2}$. Dropping the arguments in $u_{i,t}(\sigma_{i,t}, \sigma_{-i,t})$, we define agent i's information set $G_{i,t}(k)$ as

$$G_{i,t}(k) = \{(u_{j,t}, \sigma_{j,t}) \mid j = i - k, ..., i + k\}.$$

If $k = \lfloor \frac{N}{2} \rfloor$, the information set contains information about all agents on the circle. In this case, we say that agents have *global information.* If $k < \lfloor \frac{N}{2} \rfloor$ the information set contains not all relevant information. If $k = 1$, agents observe strategies and payoffs of their direct neighbors only. We call this information setting *local information.*

Now let us turn to the question how agents use information. Following Eshel et al. (1998), we assume that at the end of a period t, the agents observe $G_{i,t}(k)$. In the following period, they play the strategy that has

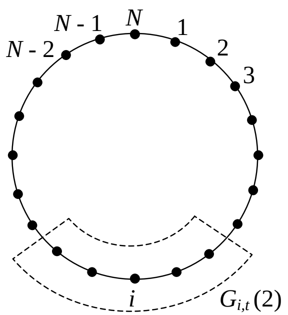


Figure 2.1: Positions and information set.

generated the highest average payoff.[6] If they observe but one strategy within their information set, they play this strategy next period.

Let $d_{j,t}$ denote an indicator variable, which takes the value 0 if agent j plays C and the value 1 if he plays D. Then, if strategies C and D are observed in $G_{i,t}(k)$, the observed payoff difference $\Delta_{i,t}$ is

$$\Delta_{i,t} = \frac{\sum_{j=i-k}^{i+k} d_{j,t} \cdot u_{j,t}}{\sum_{j=i-k}^{i+k} d_{j,t}} - \frac{\sum_{j=i-k}^{i+k} (1 - d_{j,t}) \cdot u_{j,t}}{\sum_{j=i-k}^{i+k} (1 - d_{j,t})}. \tag{2.2}$$

The first term is the average payoff of those agents in the information set that play D and the second term is the average payoff of those agents that play C. The imitation dynamics satisfies the following rule.

Definition 1 *The imitation rule is*

$$\sigma_{i,t+1} = \begin{cases} \sigma_{i,t} & \text{if} \quad \Delta_{i,t} = 0 \\ C & \text{if} \quad \Delta_{i,t} < 0 \\ D & \text{if} \quad \Delta_{i,t} > 0 \end{cases} . \tag{2.3}$$

if both strategies C and D are observed in $G_{i,t}(k)$. Otherwise, the agent continues to use the strategy played in the prior period.

[6] Agents who choose strategies according to an imitation rule can interpret the information they receive in two different ways. They can either imitate the most successful player or the most successful strategy they observe. The former imitation rule is used by Vega-Redondo (1997) and Alos-Ferrer et al. (2000) and others, the latter e.g. by Ellison and Fudenberg (1995). We will adhere to an imitation rule of the second kind, where success of a strategy is measured by its average payoff.

The imitation rule implies that if $\Delta_{i,t} = 0$, then $\sigma_{i,t+1} = \sigma_{i,t}$. For $\Delta_{i,t} \lesseqgtr 0$, we have $\sigma_{i,t+1} = C$, respectively D. If a strategy is not observed in an agent's information set, then the agent continues to use the strategy of the current period. Note that an agent's behavior depends solely on the strategies observed in his neighborhood in the immediate past. That is, neither the shadow of the future nor the shadow of the more distant past bear any weight for the choice of strategy (Berninghaus et al. 2003).

In the following we will suppress the time index t. By y we denote the number of agents playing C. This allows us to write agent i's payoff of playing $d_i \in \{0, 1\}$ as

$$u_i(d_i, y) = \frac{y-1}{N-1}\frac{w}{2} + d_i\left(\frac{N}{N-1}\frac{w}{2} - c\right). \tag{2.4}$$

The first term in (2.4) is the period payoff of an agent who plays C ($d_i = 0$). The second term, which depends only on the population size N and the cost c, is the additional payoff for an agent of playing D ($d_i = 1$). For a homogeneous population, (2.4) implies that an agent i who observes C and D in $G_i(k)$ compares $u_j(1, y)$ and $u_h(0, y)$ for some j, h. Thus, $\Delta_i = u_j(1, y) - u_h(0, y) = \frac{N}{(N-1)}\frac{w}{2} - c$. Note that Δ_i is independent of y.

The model exhibits a finite population effect. Because agents do not play against themselves, Δ_i only approaches the value $\frac{w}{2} - c$ as N goes to infinity. Thus, for any finite N and for $c \in \left(\frac{w}{2}, \frac{N}{N-1}\frac{w}{2}\right)$, D is a strictly dominated strategy. Nevertheless $\Delta_i > 0$ for all i who observe the strategy D in their neighborhood. That is, despite D being a strictly dominated strategy, all individuals will end up playing it. The reason for this result is that an agent playing D is matched with $N - y - 1$ players that are cheating too, while an agent playing C is matched with $N - y$ agents playing D. This increases the benefit of using strategy D relative to C, such that a strictly dominated strategy is played.

2.3.3 Absorbing States

A state is a specification of which agents play C and which play D. At time t, we describe the state s_t of the system by an N-tuple

$$s_t = (\sigma_{1t}, \sigma_{2t}, ..., \sigma_{Nt}) \in S \equiv \{C, D\}^N,$$

where S is the set of possible states. If i and j are two possible strategy states in S, p_{ij} is the probability that the imitation rule changes the system to state j given that i is the current state. The imitation rule in (2.3) and the non-stochastic nature of the payoffs result in a deterministic process such that p_{ij} is either 0 or 1. The collection $\{p_{ij}\}_{i,j\in S}$, together with an initial state, is a Markov process on S. We will refer to this Markov process as the imitation dynamics of our model.

We are interested in the absorbing states of the imitation dynamics, which are defined as in Eshel et al. (1998).

Definition 2 *A set of states is absorbing if it is a minimal set of states with the property that the Markov process can lead into this set but not out of it.*

An absorbing set of states may contain only one state. If an absorbing set contains more than one state, the Markov process cycles between the states contained in the absorbing set.

From now on, we normalize $w = 1$ (and consequently c is now assumed to be $c \in (0, 1)$). Moreover, we concentrate on the polar cases; i.e., either the size of the information set is $k = 1$ or $k = \lfloor \frac{N}{2} \rfloor$. In the case of local information ($k = 1$), each agent observes the strategies and the payoffs of his direct neighbors only. In the case of global information ($k = \lfloor \frac{N}{2} \rfloor$), each agent obtains information about the strategies and the payoffs of all agents on the circle.

2.3.4 The Role of Information

As was discussed in Section 2.2.3, throughout this chapter we study the spread of D in a population which is characterized by the feature that in $t = 1$ all agents but one play C.

Global Information

If $\sigma_i = C$ (D) $\forall i$, we denote this state by $\vec{C}$ $(\vec{D})$. Under global information, we obtain the following result.

Proposition 1 *Suppose agents are homogenous and have global information. If* $c \lesseqgtr \frac{N}{2(N-1)}$, *the absorbing state* $\vec{D}$, *respectively* $\vec{C}$, *is reached in period* 2.

The proof is straightforward. If all agents observe the strategies and payoffs of all other agents, Δ_i is identical for all $i = 1, 2, ..., N$. Consequently, when an agent introduces D in period 1, depending on whether Δ_i is positive or negative, all agents will play C or D from period 2 on until the end of time. This result has been shown to hold by Kandori et al. (1993) in a more general setting.

Local information

We now consider the case where the agents observe their immediate neighbors only ($k = 1$).

Proposition 2 *Suppose agents are homogenous and have local information. If* $c < \frac{N}{2(N-1)}$, *the absorbing state* $\vec{D}$ *is reached in period* $t = 1 + \lfloor \frac{N}{2} \rfloor$. *If* $c > \frac{N}{2(N-1)}$, *the absorbing state* $\vec{C}$ *is reached in period* 2.

By intuition the proof of Proposition 2 is established as follows. If $c < \frac{N}{2(N-1)}$, Δ_i is positive and all agents that are aware of strategy D imitate it in the following period. Because $k = 1$, it takes $\lfloor \frac{N}{2} \rfloor$ periods until all agents have learned and adopted D. If $c > \frac{N}{2(N-1)}$, Δ_i is negative. Consequently, D dies out immediately.

Comparing Proposition 1 and Proposition 2 makes obvious that the size of the information set only affects the time elapsing until the absorbing state is reached. In particular, it does not matter which player introduces strategy D because all agents are identical. This motivates to introduce heterogeneity among the agents to see how this influences the imitation dynamics.

2.4 Heterogenous Agents with Global Information

In this section we investigate the role of heterogeneity when agents are globally informed. Local information is then analyzed in Section 2.5.

2.4.1 Asymmetric Games

Heterogeneity of agents is introduced by assuming that agents are either of high type (H) or low type (L). An agent's type is neither known by himself nor by any other agent. As before, agents are matched pairwise and the payoffs are as follows:

If an H-type is matched to an L-type and both agents use the same strategy, the H-type wins with certainty.

If an H-type is matched to an L-type and only the L-type uses D, then the L-type wins with certainty. Thus, an L-type prevails over an H-type if and only if he plays D and the latter plays C.

The payoff matrices for the asymmetric matches are

$$A_{H,L} = \begin{pmatrix} w & 0 \\ w-c & w-c \end{pmatrix} \quad \text{and} \quad A_{L,H} = \begin{pmatrix} 0 & 0 \\ w-c & -c \end{pmatrix},$$

where, for example, $A_{H,L}$ denotes the payoffs to agent of type H when playing against an agent of type L. For symmetric matches, the matrices are $A_{H,H} = A_{L,L} = A$ as in (2.1).

If two agents of the same type meet, the stage game is a prisoner's dilemma with D as the dominant strategy if $c < 1/2$. This is similar as with homogenous agents. But in contrast to that, if two agents of opposite type meet the game is asymmetric and has a unique Nash equilibrium in mixed strategies. In this case, the stage game is a matching penny game where each player's best response is the strategy not chosen by the other agent.

We denote by y_H the number of H-types playing C, and by y_L the number of L-types playing C. The numbers of H-types and L-types prevailing in a population are denoted by n_H and n_L, respectively, with $n_H + n_L = N$. The period payoffs of H-types and L-types are

$$u_H(d_i, y_H, y_L) = \frac{y_H + 2y_L - 1}{2(n_H + n_L - 1)} + d_i \left(\frac{n_H + 2n_L - 2y_L}{2(n_H + n_L - 1)} - c \right) \quad (2.5)$$

$$u_L(d_i, y_H, y_L) = \frac{y_L - 1}{2(n_H + n_L - 1)} + d_i \left(\frac{n_L + 2y_H}{2(n_H + n_L - 1)} - c \right). \quad (2.6)$$

Note that the second term in (2.5) depends negatively on y_L, while the second term in (2.6) depends positively on y_H. The additional value of playing D for H-types (i.e. the second term in (2.5)) decreases with the

number of L-types playing C. The additional value of playing D for L-types increases with the number of H-types playing C.

This is quite intuitive. Since an H-type prevails over an L-type playing C with certainty, playing D becomes less attractive the more L-types play C. On the other hand, because an L-type prevails over an H-type if and only if he plays D and the H-type C, D becomes more attractive to L-types as the number of H-types playing C increases.

2.4.2 The Role of Information

The imitation rule (2.3) still applies and we continue to study the spread of D from an initial situation where all agents but one play C. A state where all agents of the same type play the same strategy is denoted by $\overrightarrow{\sigma_H}\overrightarrow{\sigma_L}$ where the first component means that all H-types play σ_H, the second component that all L-types play σ_L.

Proposition 3 *Suppose agents are heterogenous and have global information. If* $c \lessgtr \frac{n_H+n_L}{2(n_H+n_L-1)}$, *the absorbing state* $\vec{D}\vec{D}$ ($\vec{C}\vec{C}$) *is reached in period* 2.

Proposition 3 states that if information is global, it is immaterial which type of player innovates D. Intuitively, this sounds convincing. With global information all agents have the same information. Consequently, all agents follow the same decision rule.

From Propositions 2 and 3 follows that heterogeneity does not affect the absorbing states, if agents have information about strategies and payoffs of all agents but no information about types. However, as we will see, heterogeneity matters either if agents can recognize the types of all agents, or if they are locally informed only about strategies and payoffs.

Before we consider local information in Section 2.5, let us consider the model for the case that agents observe all strategies, all payoffs, and all types. We call this information structure the "complete information" benchmark. With complete information, agents of the same type make the same strategy decisions. Note that for the complete information benchmark we interpret the imitation rule (2.3) as follows. When applying (2.3), agents of the same type compare only payoffs and strategies across agents of their own type. [7] Based upon this, we obtain the following result.

Proposition 4 *Suppose agents are heterogenous and have complete information. If* $c < \min\{\frac{n_H}{2(N-1)}, \frac{n_L}{2(N-1)}\}$*, the absorbing state is* $\vec{D}\vec{D}$*. If*

$$c \in \left[\min\{\frac{n_H}{2(N-1)}, \frac{n_L}{2(N-1)}\}, \max\{\frac{n_H + 2n_L}{2(N-1)}, \frac{2n_H + n_L}{2(N-1)}\}\right],$$

the absorbing set is $\left\{\vec{C}\vec{D}, \vec{D}\vec{D}, \vec{D}\vec{C}, \vec{C}\vec{C}\right\}$*. If* $c > \max\{\frac{n_H+2n_L}{2(N-1)}, \frac{2n_H+n_L}{2(N-1)}\}$*, the absorbing state is* $\vec{C}\vec{C}$*.*

Since all players have both strategies available at any point in time, agents of the same type will always play the same strategy. Obviously, if costs are small (large), all agents play D (C). However, in contrast to the game with global information, there is an absorbing set in which agents cycle between C and D. This absorbing set is attained if costs are such that D (C) pays for a single agent of either type when all others play C (D).[8]

2.5 Heterogenous Agents with Local Information

With heterogeneity among agents and local information, the allocation of types along the circle matters because it affects the payoffs of the strategies C and D which an agent observes in his information set $G_i(1)$. Consequently, in contrast to the case with homogeneous agents, $G_i(1)$ does not reveal the true benefit of a strategy to a player. For example, a large payoff of a neighbor can now be due to either the strategy chosen (which is observed) or the unobservable H-type. In this section, we first classify the agents according to their location, which determines the perceived period

[7] If no agent of one type plays D (C), we assume that agents of this type can calculate the hypothetical payoff of playing D (C) for their type and have both strategies at their disposal at any point in time. Otherwise, D would trivially at most spread among the types where innovation occurred.

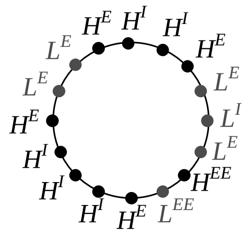
[8] Cycling also occurs in Berentsen and Lengwiler (2004). In this model the authors consider the replicator dynamics in a model with heterogenous agents. The stage game is a Prisoner's Dilemma when two agents of the same type meet or a matching penny game when two agents of opposite type meet.

payoff of the strategies C and D. We then investigate the implication of local information and heterogeneity for the imitation dynamics.

2.5.1 The Role of Information

We introduce the following notation in order to classify the agents according to their location on the circle:

An agent whose immediate neighbors are of the same type is called an *interior player*. We denote interior players by H^I and L^I, respectively. An agent who has an H-type as a neighbor on one side and an L-type on the other side is called an *edge player*. We denote edge players by H^E or L^E, respectively. A further case has to be defined: the *double-edge players*. An H-type whose two neighbors are L-types, is a double-edge player and abbreviated by H^{EE}. Accordingly, an L^{EE} is an L-type who is located such that both his neighbors are H-types. For the purpose of clarification, we display an example:



In order to simplify the display of populations, we will write populations and fragments of populations as a sequences of types. Consequently, the above example can be depicted as

$$H^I H^E L^E L^I L^E H^{EE} L^{EE} H^E H^I H^I H^I H^E L^E L^E H^E H^I \, .$$

Note that the players at both ends (H^I on the left and H^I on the right) are immediate neighbors on the circle.

Let us now determine the strategy choices for the three locations by calculating the decision terms Δ introduced above.

Interior Player An interior player has only information about agents of his own type. Consequently, an interior player only observes differences in payoffs if different strategies have been played. We write these payoff differences as

$$\Delta_{H^I}(y_H, y_L) = u_H(1, y_H, y_L) - u_H(0, y_H, y_L) = \frac{n_H + 2n_L - 2y_L}{2(N-1)} - c \quad (2.7)$$

$$\Delta_{L^I}(y_H, y_L) = u_L(1, y_H, y_L) - u_L(0, y_H, y_L) = \frac{n_L + 2y_H}{2(N-1)} - c\,. \quad (2.8)$$

If $\Delta_{H^I} \lesseqgtr 0$ (or $\Delta_{L^I} \lesseqgtr 0$) , then H^I (L^I) plays C, respectively D next period.

Edge Player Let us first consider an edge player of type H. Such a player has an L-type and an H-type as neighbors, i.e. either LH^EH or HH^EL is the respective sequence on the circle. We concentrate on LH^EH because HH^EL can be analyzed in the same way. An H^E-player faces eight ($= 2^3$) possible strategy strings in $G_{i,t}(1)$. Two strings are CCC and DDD. In this case the agent does not change his strategy. The other six strings are

$$\underbrace{CDD}_{1^{st}} \quad \underbrace{CCD \quad CDC}_{2^{nd}} \quad \underbrace{DDC \quad DCD}_{3^{rd}} \quad \underbrace{DCC}_{4^{th}}\,. \quad (2.9)$$

Consider, for example, the first term in (2.9). It means that "L plays C, H^E plays D and H plays D." The strategy strings CCD and CDC do not differ with respect to the observed average payoffs. In either case, the L-type and one H-type play C, while the other H-type plays D. Similarly, the strategy strings DDC and DCD yield also the same average payoffs for D and C, respectively.

We can summarize the decisions of H^E by considering the observed differences $\Delta^q_{H^E}$, where the superscript q refers to the rank of the term in (2.9). In the Appendix we show that the ranking orders as follows:

$$\Delta^1_{H^E} \geq \Delta^2_{H^E} \geq \Delta_{H^I} \geq \Delta^3_{H^E} \geq \Delta^4_{H^E} \quad (2.10)$$

Thus, if for a clean edge player in a DCC string $\Delta^4_{H^E} > 0$, all edge players and all interior players will play D in the following period.

An edge player of type L has an H-type and an L-type in his information group. We analyze HL^EL because LL^EH can be analyzed accordingly. The possible strategy strings in the information group of L^E are given in (2.9). As is shown in the Appendix

$$\Delta^4_{L^E} \geq \Delta^3_{L^E} \geq \Delta_{L^I} \geq \Delta^2_{L^E} \geq \Delta^1_{L^E}\,. \quad (2.11)$$

Next we show that some agents systematically over- or underestimate the true benefit of the illegal strategy D. Recall that an interior player only observes agents of his own type. Consequently, his decision term Δ_{L^I} (or Δ_{H^I}) reflects the true payoff difference of the two strategies for his type, i.e. an interior player's assessment of a strategy is not distorted by heterogeneity. However, according to (2.10) and (2.11), edge players systematically over- or underestimate the payoffs of C or D to their type as defined below.

Definition 3 *An agent overestimates (underestimates) the payoff of strategy D for his type if his decision term Δ is greater (smaller) than the decision term of an interior player of his type.*

For example, edge players of type H overestimate D whenever their L-type neighbor plays C, i.e. $\Delta^1_{H^E} \geq \Delta^2_{H^E} \geq \Delta_{H^I}$. The reason for this is that the payoff of an L-type using C is always zero except when matched to another L-type player using C. In the later case his payoff is $1/2$. In contrast, an H-type who plays C receives a positive payoff when matched to another H-type or to an L-type using C. Consequently, an edge player of type H underestimates the benefit of strategy C, respectively overestimates D, when his L-type neighbor plays C.

To recapitulate, there are two crucial features of local information. First, certain agents (edge players and double-edge players) under- or overestimate strategy D, respectively, C. There is no such effect with *complete* information where each player type knows the true benefit of each strategy for his type.

Second, local information permits some agents (interior players) to observe the true payoff difference of the two strategies for his type as explained above. In contrast, with *global* information and heterogenous players no agent ever observes the true payoff difference for his type.

Double-Edge Player Finally, let us investigate the behavior of double-edge players. The information group for an H^{EE}-type and an L^{EE}-type respectively are composed of the following types:

$$LH^{EE}L \quad \text{and} \quad HL^{EE}H.$$

Thus, the relevant strategy strings are

$$\underbrace{DCD}_{1^{st}} \quad \underbrace{CCD \quad DCC}_{2^{nd}} \quad \underbrace{DDC \quad CDD}_{3^{rd}} \quad \underbrace{CDC}_{4^{th}}\,. \quad (2.12)$$

In the Appendix we show that the following ranking holds,

$$\Delta^3_{HEE} > \Delta_{H^I} > \Delta^2_{HEE} \quad \text{and} \quad \Delta^2_{LEE} > \Delta_{L^I} > \Delta^3_{LEE}. \tag{2.13}$$

Like edge players, double-edge players over- or underestimate the payoff of strategy D to their type.

2.5.2 Maximal Segregation

Recall the main question of this analysis: How does strategy D evolve in a population in which initially all players but one play C? Two factors determine the spread of D: The allocation of types along the circle on the one hand, and the location and type of the innovator on the other.

In the following we consider a distribution of types that we call maximal segregation. In such a population H-types and L-types are allocated in two clusters as follows:

$$HHH...HHLLL...LLL\,.$$

In a maximally segregated population, there are only two edge players for each type and no double-edge players. In order to simplify the analysis, we assume that $n_H = n_L = n = \frac{N}{2}$.

Innovation and absorbing states

As explained above, for each type there are three classes of agents; interior, edge and double-edge players. Within the same class agents may choose different strategies because they have different information sets. Consequently, we have to distinguish the *location of the innovator*. Innovation by an interior player has a different implication for the imitation dynamics than an innovation by an edge player. Moreover, we have also to distinguish among interior players. An innovation through an interior player who is located within other interior players has different consequences than an innovation from an interior player who is located next to an edge player. We call these special interior players *next-to-edge players* and give them the superscript NE, while we still denote all other interior players by superscript I.

Proposition 5 *The innovator can either be an L-type or an H-type.*
(I) Suppose the innovator is an L-type:

For $c < \frac{2n}{2(N-1)}$ the absorbing state is $\vec{D}\vec{D}$, for $c \in \left(\frac{2n}{2(N-1)}, \frac{3n}{2(N-1)}\right)$ it is $\vec{C}\vec{D}^*$, and for $c > \frac{3n}{2(N-1)}$ it is $\vec{C}\vec{C}$.
(II) Suppose the innovator is an H-type:

(i) If he is an interior player, then for $c \lesseqgtr \frac{n}{2(N-1)}$, the absorbing state is $\vec{D}\vec{D}$, respectively $\vec{C}\vec{C}$.

(ii) If he is a next-to-edge player, then for $c < \frac{n+1}{2(N-1)}$, the absorbing state is $\vec{D}\vec{D}$, for $c \in \left(\frac{n+1}{2(N-1)}, \frac{n+2}{2(N-1)}\right)$, it is $\vec{D}\vec{C}^*$, for $c \in \left(\frac{n+2}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$ it is $\vec{D}\vec{D}$, and for $c > \frac{2n-\frac{1}{2}}{2(N-1)}$, it is $\vec{C}\vec{C}$.

(iii) If he is an edge player, then for $c < \frac{n+1}{2(N-1)}$, the absorbing state is $\vec{D}\vec{D}$, for $c \in \left(\frac{n+1}{2(N-1)}, \frac{n+2}{2(N-1)}\right)$, it is $\vec{D}\vec{C}^*$, for $c \in \left(\frac{n+2}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$ it is $\vec{D}\vec{D}$, for $c \in \left(\frac{2n-\frac{1}{2}}{2(N-1)}, \frac{3n-1}{2(N-1)}\right)$, it is $\vec{C}\vec{D}^*$, and for $c > \frac{3n-1}{2(N-1)}$, it is $\vec{C}\vec{C}$.

The absorbing state $\vec{C}\vec{D}^*$ ($\vec{D}\vec{C}^*$) is identical to $\vec{C}\vec{D}$ ($\vec{D}\vec{C}$) except that edge players of type L play C or cycle between C and D depending on c and n. Interestingly, the location of the innovator is irrelevant if the innovator is an L-type. However, if the innovator is an H-type, the location matters. Finally, the number of agents playing C can decrease in c when the innovation occurs through an H^N- or an H^E-type.

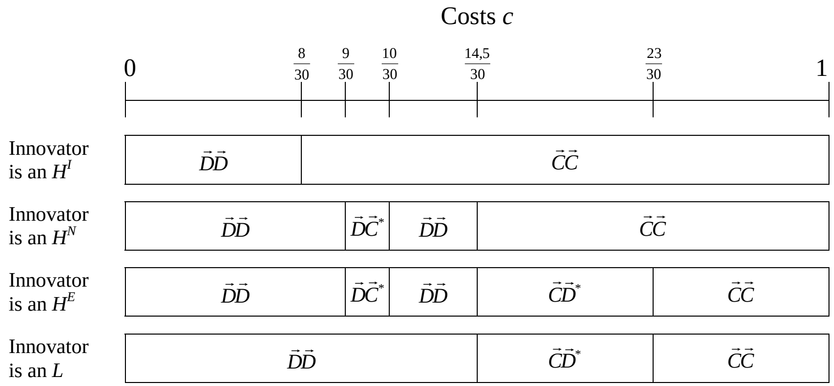


Figure 2.2: Absorbing states in a maximally segregated, finite population.

Let us illustrate Proposition 5 for $n = 8$. As Figure 2.2 shows, if the cost c is drawn at random, then the population is more likely to end up in the absorbing state $\vec{D}\vec{D}$ when the innovator is an L-type than when he is an H-type. Furthermore, Figure 2.2 illustrates, that among the H-types the location of the innovator is crucial. If the innovation arises from an H^N- or an H^E-type, then again $\vec{D}\vec{D}$ is more likely than if the innovation arises from an H^I-type. In this sense, innovations by edge or next-to-edge H-types have similar consequences for the imitation dynamics as those by L-types. Finally, Figure 2.2 shows that the number of agents playing C is non-monotonic in c.

An interesting case is the limiting case when there is no finite population effect ($n \to \infty$), as depicted in Figure 2.3. In this case the absorbing state

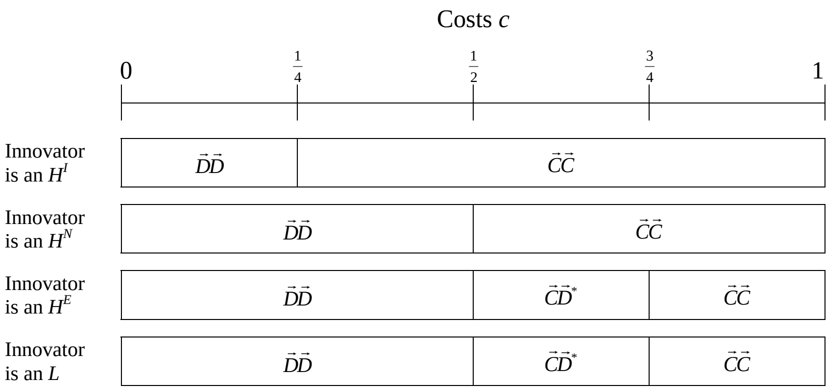


Figure 2.3: Absorbing states in a maximally segregated, infinite population.

$\vec{D}\vec{C}^*$ disappears. Indeed, without a finite population effect, the difference between an innovation by an H^I-type and a H^N- or H^E-type can be seen more easily. For $c \in (1/4, 1/2)$ innovating H^E- and H^N-types infect L-type players, which does not happen if the innovator is an H^I-type. Note that in this case innovations by L-types or H^E- and H^N-types have almost the same consequences for the imitation dynamics.

2.5.3 Minimal Segregation of Types

After having characterized the absorbing states when the population is maximally segregated, we now consider the polar case of a minimally segregated population. This means that we look at a population in which types are located as follows:

$$HLHLHLHL...HLHLHLHL..$$

Evidently, a minimally segregated population consists of double-edge players only. Consequently, there are only two different positions where the strategy D can be introduced. These are H^{EE} and L^{EE}, respectively. Again, let us assume $n_H = n_L = n$.

Absorbing States

Recall that in a maximally segregated population, all agents of the same type play the same strategy in the absorbing state (with the only exception of edge-players of type L in some absorbing states, see Proposition 5). In contrast to that, in a minimally segregated population, agents of the same type will not necessarily end up playing the same strategy. For analyzing this, we have to introduce some additional notation.

We denote by $\tilde{C}_{y_H}\tilde{C}_{y_L}$ a strategy state where the number of H-types and L-types playing C is y_H and y_L, respectively. In such a state, all agents of either type who play the same strategy are next to each other. Let $\bar{y}$ be the greatest nonnegative odd (even) integer smaller than

$$2c\,(N-1) - n - \frac{1}{2}$$

if n is even (odd). Note that if $2c\,(N-1) - n - \frac{1}{2} < 0$, i.e. $c < \frac{n+\frac{1}{2}}{2(N-1)}$, then $\bar{y} = 0$. In this case, the absorbing state is $\tilde{C}_0\tilde{C}_0$ where all agents play D. It is again possible that an absorbing set is attained in which two L-types cycle between D and C. We denote such an absorbing set by $\tilde{C}_{y_H}\tilde{C}^*_{y_L}$.

Proposition 6 *The innovator can either be an L-type or an H-type.*
(I) Suppose the innovator is an H-type:
For $c < \frac{n+\frac{1}{2}}{2(N-1)}$, the absorbing state is $\tilde{C}_0\tilde{C}_0$, for $\frac{n+\frac{1}{2}}{2(N-1)} < c < \frac{2n-\frac{1}{2}}{2(N-1)}$ it is $\tilde{C}_{\bar{y}}\tilde{C}^*_{\bar{y}-1}$, for $\frac{2n-\frac{1}{2}}{2(N-1)} < c < \frac{3n-1}{2(N-1)}$ it is $\tilde{C}_{n-1}\tilde{C}_n$, and for $c > \frac{3n-1}{2(N-1)}$, it is

$\tilde{C}_n\tilde{C}_n$.

(II) Suppose the innovator is an L-type:

For $c < \frac{n+\frac{1}{2}}{2(N-1)}$, *the absorbing state is* $\tilde{C}_0\tilde{C}_0$, *for* $\frac{n+\frac{1}{2}}{2(N-1)} < c < \frac{2n+\frac{1}{2}}{2(N-1)}$ *it is* $\tilde{C}_{\bar{y}}\tilde{C}^*_{\bar{y}-1}$, *and for* $c > \frac{2n+\frac{1}{2}}{2(N-1)}$ *it is* $\tilde{C}_n\tilde{C}_n$.

Several comments seem to be helpful for a better understanding of the results. First, in a minimally segregated population the type of the innovator does not affect the absorbing states significantly. It only matters if $\frac{2n-\frac{1}{2}}{2(N-1)} < c < \frac{3n-1}{2(N-1)}$. In this case, the absorbing state is $\tilde{C}_{n-1}\tilde{C}_n$ if the innovator is an H-type.

Second, if the innovator infects its neighbors such that strategy D begins to spread, the spread can be only blocked by H-types. Consequently, in any absorbing state where both strategies survive and where more than one player adopts strategy D there will be always one $L-$player more using D than H-players, i.e. the absorbing state is of type $\tilde{C}_{\bar{y}}\tilde{C}^*_{\bar{y}-1}$.

Third, the absorbing state $\tilde{C}_{n-1}\tilde{C}_n$ is special because the innovator is a $H-$type, which is not able to infect his L-type neighbors. Nevertheless, he continues to use D because strategy D yields a higher payoff in his information set. Consequently, the initial strategy string is stationary.

We illustrate Proposition 6 for $n = 8$ in Figure 2.4.

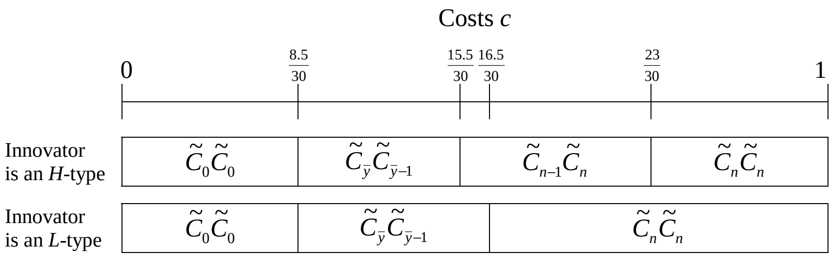


Figure 2.4: Absorbing states in a minimally segregated, finite population.

Figure 2.4 shows that the absorbing states are independent of the innovator's type for costs smaller than $\frac{2n-\frac{1}{2}}{2(N-1)}$ and costs higher than $\frac{3n-1}{2(N-1)}$. For the remainder of the cost interval, the absorbing states for an H-type innovator can maximally differ by the strategy choice of two agents from the absorbing states for an L-type innovator.

Finally, we also consider the limiting case without finite population effect ($n \to \infty$). With n going to infinity, $\bar{y}$ goes to infinity too. So we cannot indicate the respective absorbing state with $\bar{y}$. We use the share of agents playing C in dependence of the costs c instead. Consequently, an absorbing state $\tilde{C}_{4c-1}\tilde{C}_{4c-1}$ means that for $c = 0.3$ one fifth $(= 4 * 0.3 - 1)$ of the H-types and one fifth of the L-types play C.

From Figure 2.5 we can see that without finite population effect the type of innovator is irrelevant. If $c < 1/4$, then all players will end up using D.

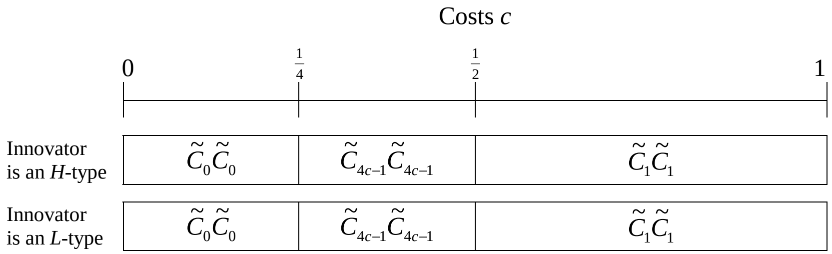


Figure 2.5: Absorbing states in a minimally segregated, finite population.

For $1/4 < c < 1/2$, the fraction of H-types and L-types using C is equal and strictly increasing in c. Finally, for $c > 1/2$, all agents play C. This is interesting because with global information in the limiting case when the number of players is large we get the result that if $c \lessgtr \frac{1}{2}$, the absorbing state is $\vec{D}\vec{D}$ ($\vec{C}\vec{C}$) (see Proposition 3). Thus, local information reduces the spread of the strategy D if agents are minimally segregated.

2.5.4 Maximal versus Minimal Segregation

Finally, let us compare the absorbing states of a maximally and of a minimally segregated population. To this end, we calculate the expected share of agents playing C in the absorbing state when each agent is equally likely to innovate D. We focus on large populations ($n \to \infty$) such that the finite population effects can be neglected. Another consequence of this assumption is that the role of edge players (of which there are but two of each type in the maximally segregated population and none in the minimally segregated population) becomes negligible.

Figure 2.5.4 depicts the expected shares of agents playing C in the absorbing states for maximally and minimally segregated populations. The left-hand panel displays the shares for both populations when D is introduced by an H-type. The right-hand panel displays these shares if innovation occurs by an L-type.

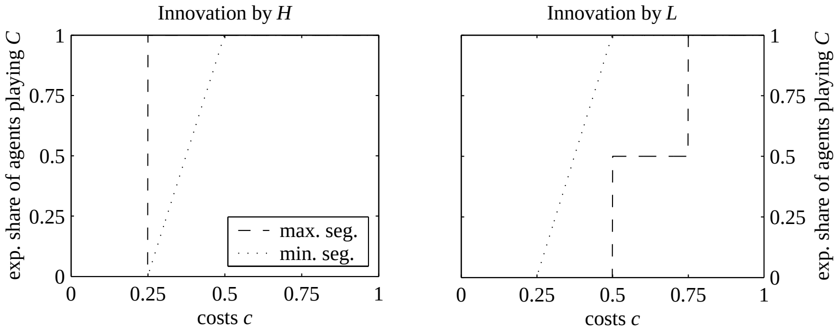


Figure 2.6: Comparing maximal and minimal segregation.

There are two observations worth pointing out. First, in a minimally segregated population the location of the innovator does not matter. Consequently, for a minimally segregated population the curves depicting the share of agents playing C (the dotted curves in Figure 2.5.4) are identical in both panels.

Second, if the innovator is a low type, then a minimally segregated population is more resistant against the illegal strategy D than a maximally segregated population: The curve of the maximally segregated population lies more to the right than the one of the minimally segregated population. This means that in a maximally segregated population the costs leading to a higher share of C-playing agents are higher than in a minimally segregated population. The reason for this result is that in a minimally segregated population each low type (including the innovator) is surrounded by two high types who are less prone to imitate D. In contrast, if the innovator is a high type, then a maximally segregated population is more resistant against D than a minimally segregated population. The reason for this result is that in a maximally segregated population the innovator is again surrounded by high types which are less prone to imitate D.

Proposition 7 *Consider a large population ($n \to \infty$) and suppose that each agent is equally likely to innovate the illegal strategy D.*
If $c < \frac{2}{8}$, all agents play D in the absorbing state for both distributions of types.
If $\frac{2}{8} < c < \frac{3}{8}$, more agents play D in a minimally segregated population than in a maximally segregated one.
If $\frac{3}{8} < c < \frac{6}{8}$, less agents play D in a minimally segregated population than in a maximally segregated one.
Finally, if $c > \frac{6}{8}$, all agents play C in the absorbing state for both distributions of types.

These results immediately follow from Propositions 5 and 6.

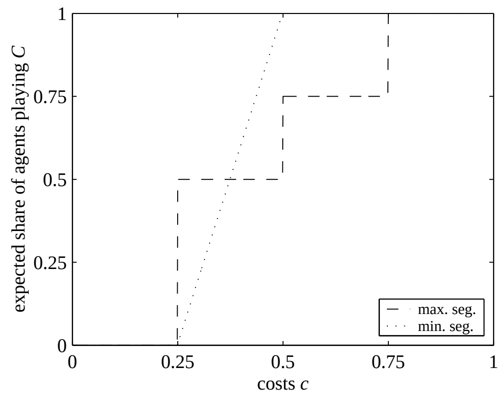


Figure 2.7: Comparing maximal and minimal segregation, one panel.

As Figure 2.7 shows, for $c \in (\frac{3}{8}, \frac{6}{8})$ the minimally segregated population exhibits a higher share of agents playing C in the absorbing state than the maximally segregated population. The reason is that being located between L-types, the H-types are able to block the spread of strategy D. However, for sufficiently low cheating costs, i.e. for $c \in (\frac{2}{8}, \frac{3}{8})$, the maximally segregated population exhibits a higher share of agents playing C, simply since in a maximally segregated population L-types never observe D if the innovator is a H-type. However, in a minimally segregated population, if D is introduced by a H-type, two L-type will observe it. And since L-types are more likely to be infected than H-types, the population share of agents playing D is larger with minimal segregation.

2.6 Extensions

Of course our analysis so far might be extended into several directions. Two of them will be discussed at some detail.

In Section 2.6.1 we first consider mutations. So far we have studied how an illegal activity spreads in a fair playing population, if only a single agent discovers it. We have neither considered that there could be more than just one innovating agent at the same time nor that innovations could take place after the first period. By considering mutation, we now allow for random innovations in all periods. How robust are our results with respect to random innovations? Can we expect similar results if an illegal strategy appears randomly in a population?

In Section 2.6.2 we consider gradual heterogeneity and gradual effectiveness of D. Our results so far have been based on the assumption that a cheating agent always wins against a fair playing agent, and that a high type always prevails over a low type if both play the same strategy. Both are strong assumption which we intend to generalize in Section 2.6.2. We sketch how the absorbing states change when D is less than perfectly effective or when H-types do not win with certainty over L-types when both types use the same strategy. Our previous results are confirmed to hold when we give up the two assumptions.

2.6.1 Mutations

How can we introduce mutations? One option is to suppose that in each period, after imitations have occurred, an agent's strategy changes with a small probability ε.[9] As mentioned before, the imitation dynamics is a Markov chain evolving over the strategy space S. A probability distribution over S in time t is represented as a row vector ν which is an element of the 2^N-dimensional simplex. The simplex Σ_N is the set

$$\Sigma_N = \left\{ \nu \in \mathbb{R}^{2^N} \;\middle|\; \nu_i \geq 0 \text{ and} \sum_i \nu_i = 1 \text{ for } i = 1, 2, ..., 2^N \right\}.$$

[9]See Kandori et al. (1993, p.38) or Ellison (1993, p.1050) for an interpretation of ε.

The process evolves according to $\nu_{t+1} = \nu_t P$, where P is the transition probability matrix as defined in Section 2.3. Now since strategies change with probability ε after imitation, the transition probability p_{ij} is positive for all i and j, i.e. the Markov chain is regular. Thus, there exists a unique probability distribution $\mu \in \Sigma_N$ such that[10]

$$\mu P = \mu \,.$$

The vector μ is the unique stationary distribution of the regular Markov process, which does not depend on the initial probability distribution. The stationary distribution μ is stable, i.e.

$$\lim_{t \to \infty} \nu P^t = \mu \quad \forall\, \nu \in \Sigma_N .$$

From the law of large numbers for regular Markov chains we get

$$E\left[\frac{1}{T}\sum_{t=1}^{T} z_{i,t}\right] \to \mu_i \quad \text{with } z_{i,t} = \begin{cases} 1 \text{ if } s_t = i \\ 0 \text{ otherwise} \end{cases}$$

as T goes to infinity.[11] Therefore, the probabilities in the limiting distribution can be interpreted as average share of time the process spends in a given state.

The transition matrix $P(\varepsilon)$ and the stationary distribution $\mu(\varepsilon)$ depend on ε. The stationary and stable probability distribution $\mu(\varepsilon)$ describes the long-run behavior of the imitation dynamics with mutations. Since we are interested in the imitation dynamics for small ε, we consider the limiting distribution μ^*:

$$\mu^* = \lim_{\varepsilon \to 0} \mu(\varepsilon)\,.$$

The limiting distribution μ^*, if it exists, depends on the parameter values $\{c, n\}$ of our model. Even for very small populations, evaluating μ^* involves solving a large equation system of 2^{2n} variables. Instead of finding $\mu(\varepsilon)$ explicitly and taking the limit for $\varepsilon \to 0$, we approximate $\mu*$ numerically. We will describe μ^* for a maximally segregated population and the smallest population size ($n = 5$) that provides all relevant positions of innovation (i.e. H^I, H^N, H^E, and L) as described previously. Our simulations suggest the following results.

[10]See e.g. Kemeny and Snell (1960), Theorem 4.1.6.(b).

[11]See e.g. Kemeny and Snell (1960), Theorems 4.1.6.(a) and 4.2.1.

Conjecture 1 *Suppose a population is maximally segregated and the individual mutation rate ε goes to zero. If* $c < \frac{n}{2(N-1)}$, *then* $\Pr(\vec{D}\vec{D}) = 1$, *if* $\frac{n}{2(N-1)} < c < \frac{2n-2}{2(N-1)}$ *then* $\Pr(\vec{D}\vec{C}^*) = 1$, *if* $\frac{2n-2}{2(N-1)} < c < \frac{2n+2}{2(N-1)}$, *then* $\Pr(M) = 1$, *if* $\frac{2n+2}{2(N-1)} < c < \frac{3n}{2(N-1)}$ *then* $\Pr(\vec{C}\vec{D}^*) = 1$, *and if* $c > \frac{3n}{2(N-1)}$, *then* $\Pr(\vec{C}\vec{C}) = 1$.

The set M is defined as $M = \left\{\vec{D}\vec{D}, \vec{D}\vec{C}^*, \vec{D}\vec{C}, \vec{C}\vec{D}, \vec{C}\vec{D}^*, \vec{C}\vec{C}\right\}$. We illustrate our conjecture for $n = 5$ and $n \to \infty$ in the following Figures.

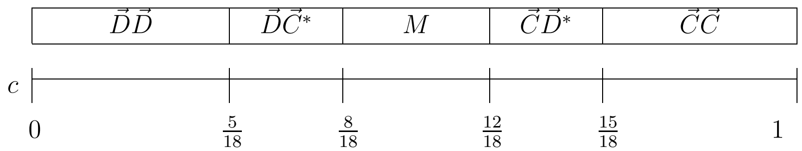


Figure 2.8: Maximal segregation with mutations, $n = 5$.

Figure 2.8 demonstrates that the absorbing state $\vec{D}\vec{D}$ arising from the costs $c \in \left(\frac{n+2}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$ in the model without now mutations disappears. Nevertheless, we still observe non-monotonicity because $\vec{D}\vec{D}$ is an element of the absorbing set M and is played with strictly positive probability.

Again, an interesting case is the limiting case when $n \to \infty$ as depicted in Figure 2.9.

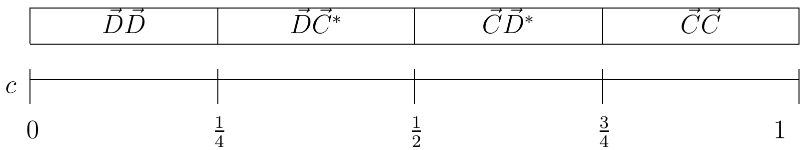


Figure 2.9: Maximal segregation with mutations, infinite n.

For $c < \frac{1}{4}$, the model with mutations and the deterministic model both give way to the absorbing state $\vec{D}\vec{D}$. If $\frac{1}{4} < c < \frac{1}{2}$, the absorbing state $\vec{D}\vec{C}^*$ is reached with probability 1 in the model with mutations. In the deterministic model, however, the absorbing state is $\vec{C}\vec{C}$ if the innovator is an H^I-type and $\vec{D}\vec{D}$ otherwise. From this we conclude that our results in the deterministic model are not robust with respect to mutations. This is

not so surprising, since due to the imitation rule, agents can only imitate strategies they observe in their information set. Thus, in the absorbing states $\vec{C}\vec{C}$ and $\vec{D}\vec{D}$, the population is locked in. This cannot occur with mutations. For $\frac{1}{2} < c < \frac{3}{4}$, the model without mutations exhibits another lock-in, which arises if the innovation occurs through an H^I- or an H^N-type. For $c > \frac{3}{4}$, the model with mutations and the deterministic model have the same absorbing state, $\vec{C}\vec{C}$.

2.6.2 Gradual Heterogeneity and Effectiveness of D

So far, we have assumed that H-types prevail with certainty over L-types if both play the same strategy. We call this complete heterogeneity. We have also used the assumption that D is perfectly effective, i.e. an L-type playing D prevails with certainty over an H-type playing C. In order to show that our results can be generalized, we now relax both assumptions.

Let p_1 denote the probability that an agent of either type playing D prevails over an agent of the same type playing C and p_2 the probability that an H-type playing C wins against an L-type playing D. The probability that an H-type wins against an L-type when both play the same strategy is denoted by p_3, while the probability that an H-type playing D wins against an L-type is denoted as p_4. With perfect effectiveness we thus had $p_1 = 1$ and $p_2 = 0$ and $p_4 = 1$, while complete heterogeneity implies $p_3 = 1$.

When relaxing the assumption of perfect effectiveness of D, we let p_1 decrease while increasing p_2 (e.g. $p_2 = 1 - p_1$). When relaxing the assumption of complete heterogeneity, we simply can decrease p_3 (see Figure 2.10). We will use a maximally segregated population to show these extensions.

What are the effects of allowing heterogeneity to be only gradual rather than complete? Our simulations are summarized in Figure 2.10. The panel on the right shows the effects of gradual heterogeneity. Gradual heterogeneity is captured by varying the parameter p_3, which is the probability that an H-type wins against an L-type when both play the same strategy.

The panel on the left shows the effects of gradual effectiveness of D. Gradual effectiveness of D means that a cheating agent does not prevail over a fair playing for sure. Therefore we vary p_1, the parameter for the probability that an agent playing D wins against an agent of his own type playing C. Both simulations are run with a population size of $N = 40$.

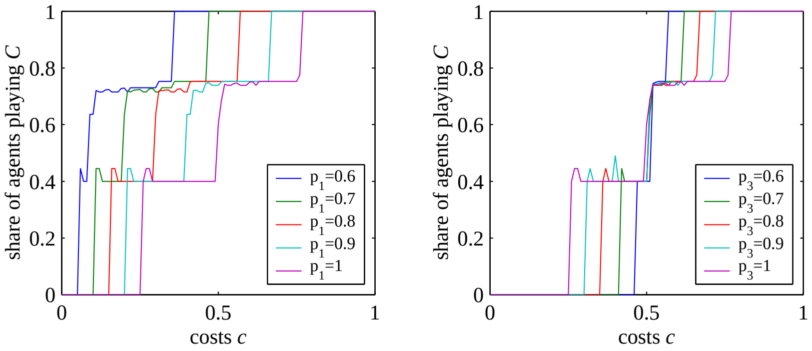


Figure 2.10: Gradual heterogeneity and gradual effectiveness of D.

The curves depicting the fraction of agents who play C are step functions in the case of complete heterogeneity, as has been shown in section 2.5.2. As heterogeneity becomes gradual, these functions become smoother and resemble more and more straight lines, as the right panel in Figure 2.10 shows. This panel depicts the fraction of agents playing C for five different values of p_3, i.e. the probability that an H-type prevails over an L-type if both play the same strategy. The resulting smoother line is not much of a surprise, given the behavior of the model with homogenous agents analyzed in Section 2.3. As the probability p_3 becomes smaller until it eventually approaches $\frac{1}{2}$, agents becomes more and more homogenous, so that the model eventually collapses to the model of Section 2.3.

The consequences of relaxing the assumption of perfect effectiveness are also easily summarized. The left panel in Figure 2.10 depicts the share of agents playing C as a function of c. Not surprisingly, as strategy D becomes less effective, the share of agents who play C increases for any value of the costs c.

We think it is fair to conclude from these two extensions that the model is robust with respect to the assumptions maintained through out the largest parts of the chapter. Qualitatively, the results do not change as we allow for less than complete heterogeneity of types and less than perfectly effective D.

2.7 Conclusions

In this chapter, we are concerned with the spread of illegal activities. Let us consider a population in which agents compete with fair means. We now assume that one agent invents an illegal activity and study the circumstances under which other agents imitate it. The innovation of this research is that we include the distortions of the information on illegal activities in our considerations: these are scarcity and non-verifiability.

In order to study these issues, we consider an evolutionary game in which agents can either use a legal or an illegal strategy. An agent playing the illegal strategy bears some direct utility cost but wins the stage game against a fair playing agent with certainty. Agents interact globally, i.e. compete against all other agents of the population. They imitate the strategy which they observed to yield the highest average payoff.

We model informational scarcity by introducing the concept of local information. With local information agents observe the payoffs received and the strategies played only of a subset of opponents. Depending on the application we may say that information cannot be hidden from all agents or that information is shared with confidants. We use a spatial model to capture this feature: Agents are allocated on a circle and under local information only neighbors observe payoffs and strategies of each other. In contrast to local information, global information implies that agents observe the strategies and the outcomes of all other agents.

The consideration of non-verifiability requires a further extension of the standard evolutionary game framework: By assuming heterogeneity of agents we introduce different payoffs for agents playing the same strategy. In an evolutionary game heterogeneity of agents is equivalent to the implementation of more than one stage game. Our evolutionary game consists of two stage games, one that defines the outcomes between agents of the same type, and one that defines the payoffs between two agents of different types. There are two types of agents where high types have a natural advantage over low types, i.e. they win the stage game with certainty if both agents apply the same strategy. We assume that the type of an agent cannot be observed by the agents. Consequently, the payoffs an agent observes for a strategy may not equal the payoff he would receive playing it. We therefore say that due to non-verifiability agents can over- or underestimate the payoff of a strategy.

Whether an agent over- or underestimates the payoff of a strategy depends on the types of agents he shares information with. It is obvious that a low type that observes the payoff of a high type overestimates the value of the strategy played by the high type. The reason is that the high types receive higher payoffs because of their natural advantage in competition. Consequently, the low type is more likely to imitate the strategy played by the high type than a strategy played by other low types. This example shows that the distribution of types matters for the absorbing states of the imitation dynamics. Naturally, the type and the location of the innovator are also decisive for the imitation dynamics.

We focus on two distribution of types: maximal (equal types are located next to each other) and minimal (agents have always neighbors of opposite type) segregation. What are the differences between the absorbing states of the two distributions?

First, the population is more likely to end up in an absorbing state where all agents act illegally if the innovator is a low type. Second, if the innovator is a low type, then a minimally segregated population is more resistant against the illegal strategy than a maximally segregated population. Third, in contrast, if the innovator is a high type, a maximally segregated population is more resistant against the illegal strategy than a minimally segregated population.

In a minimally segregated population, the location and the type of the innovator does not matter. Neither do we observe that the share of fair playing agents in an absorbing state changes significantly for small changes in the costs of the illegal activity. The reason is that high types which are less prone to imitate the illegal strategy can block the spread of the illegal strategy in every other period.

However, if a low type innovates the illegal strategy in a maximally segregated population, all low types observing it imitate it because it strengthens their position in competition. For a large number of cheating low types, high types are better off to act illegally too. Therefore a population can actually end up in an absorbing state with all agents acting illegally even for relatively high costs of the illegal activity. Consequently, we observe large changes in the frequency of the illegal activity for small changes of its costs.

By extending our model we show that these results are robust with re-

spect to the assumption on the effectiveness of the illegal strategy. Similarly, heterogeneity can be reduced gradually without changing the nature of our results.

2.A Appendix: Proofs

2.A.1 Proof of Proposition 3

If the innovating agent is an H-type, the initial numbers of clean agents are $y_H = n_H - 1$ and $y_L = n_L$. If the innovator is an L-type, these numbers are $y_H = n_H$ and $y_L = n_L - 1$. We first notice that due to global information we have that

$$\Delta_H(y_H, y_L) = \Delta_L(y_H, y_L)\,.$$

We therefore drop the type group index of the decision terms.

If the innovation arises from an H-type, the decision term for all agents is

$$\begin{aligned}\Delta(n_H-1, n_L) &= u_H(1, n_H-1, n_L)\\ &\quad - \frac{(n_H-1)\,u_H(0, n_H-1, n_L) + n_L u_L(0, n_H-1, n_L)}{N-1}\,.\end{aligned}$$

From (2.5) and (2.6) we get

$$\begin{aligned}u_H(d_i, n_H-1, n_L) &= \frac{N+n_L-2}{2(N-1)} + d_i\left(\frac{n_H}{2(N-1)} - c\right) \text{ and}\\ u_L(d_i, n_H-1, n_L) &= \frac{n_L-1}{2(N-1)} + d_i\left(\frac{N+n_H-2}{2(N-1)} - c\right), \text{ respectively.}\end{aligned}$$

Using this information we get

$$\Delta(n_H-1, n_L) = \frac{n_L+n_H}{2(n_L+n_H-1)} - c.$$

If the innovation arises from an L-type agent, the difference in the average payoffs is

$$\begin{aligned}\Delta(n_H, n_L-1) &= u_L(1, n_H, n_L-1)\\ &\quad - \frac{n_H u_H(0, n_H, n_L-1) + (n_L-1)\,u_L(0, n_H, n_L-1)}{N-1}\,.\end{aligned}$$

From (2.5) and (2.6) we get

$$\begin{aligned}u_H(d_i, n_H, n_L-1) &= \frac{N+n_L-3}{2(N-1)} + d_i\left(\frac{n_H+2}{2(N-1)} - c\right) \text{ and}\\ u_L(d_i, y_H, n_L-1) &= \frac{n_L-2}{2(N-1)} + d_i\left(\frac{N+n_H}{2(N-1)} - c\right), \text{ respectively.}\end{aligned}$$

Using this information we get

$$\Delta(n_H, n_L-1) = \frac{n_L+n_H}{2(n_L+n_H-1)} - c.$$

Note that

$$\Delta(n_H, n_L-1) = \Delta(n_H-1, n_L)\,.$$

Consequently, the origin of the innovation does not matter if agents have no type information.

2.A.2 Proof of Proposition 4

If agents have complete information, they have both strategies available at any point in time, distinguish types and are able to calculate the payoff they would have gotten having played the other strategy. Under these circumstances the imitation rule (2.3) can be interpreted as a best response dynamics. As before, we can define suitable decision terms. We have

$$\Delta_H(y_H, y_L) = u_H(1, y_H, y_L) - u_H(0, y_H, y_L) = \frac{n_H+2n_L-2y_L}{2(N-1)} - c\,, \quad (2.14)$$

$$\Delta_L(y_H, y_L) = u_L(1, y_H, y_L) - u_L(0, y_H, y_L) = \frac{n_L+2y_H}{2(N-1)} - c\,. \quad (2.15)$$

At the end of the first period the agents decide either to play

- $\vec{D}\vec{D}$, if $c < \frac{n_H}{2(N-1)}$ when H innovates (from $\Delta_H(n_H-1, n_L) > 0$) or if $c < \frac{n_H+2}{2(N-1)}$ when L innovates (from $\Delta_H(n_H, n_L-1) > 0$)
- $\vec{C}\vec{D}$, if $\frac{n_H}{2(N-1)} < c < \frac{2n_H+n_L-2}{2(N-1)}$ when H innovates (from $\Delta_H(n_H-1, n_L) < 0$ and $\Delta_L(n_H-1, n_L) > 0$) or if $\frac{n_H+2}{2(N-1)} < c < \frac{2n_H+n_L}{2(N-1)}$ when L innovates (from $\Delta_H(n_H, n_L-1) < 0$ and $\Delta_L(n_H, n_L-1) > 0$)
- $\vec{C}\vec{C}$ for $\frac{2n_H+n_L-2}{2(N-1)} < c$ when H innovates (from $\Delta_L(n_H-1, n_L) < 0$) or if $\frac{2n_H+n_L}{2(N-1)} < c$ when L innovates (from $\Delta_L(n_H, n_L-1) < 0$).

Note that $\vec{D}\vec{C}$ will not be played in the second period because of $\Delta_H(n_H - 1, n_L) < \Delta_L(n_H - 1, n_L)$ and $\Delta_H(n_H, n_L - 1) < \Delta_L(n_H, n_L - 1)$. What kind of Δ_H and Δ_L do these strategy states imply?

- All play C (situation $\vec{C}\vec{C}$), implies
$$\Delta_H(n_H, n_L) = \frac{n_H}{2(N-1)} - c \quad \text{and} \quad \Delta_L(n_H, n_L) = \frac{2n_H + n_L}{2(N-1)} - c\,,$$
- all play D (situation $\vec{D}\vec{D}$), implies
$$\Delta_H(0,0) = \frac{n_H + 2n_L}{2(N-1)} - c \quad \text{and} \quad \Delta_L(0,0) = \frac{n_L}{2(N-1)} - c\,,$$
- group H plays C and group L plays D (situation $\vec{C}\vec{D}$), implies
$$\Delta_H(n_H, 0) = \frac{n_H + 2n_L}{2(N-1)} - c \quad \text{and} \quad \Delta_L(n_H, 0) = \frac{2n_H + n_L}{2(N-1)} - c\,,$$
- group H plays D and group L plays C (situation $\vec{D}\vec{C}$), implies
$$\Delta_H(0, n_L) = \frac{n_H}{2(N-1)} - c \quad \text{and} \quad \Delta_L(0, n_L) = \frac{n_L}{2(N-1)} - c\,.$$

From (2.14) and (2.15) it is clear that $\vec{C}\vec{C}$ is an equilibrium strategy state if
$$c > \max\{\frac{n_H + 2n_L}{2(N-1)}, \frac{2n_H + n_L}{2(N-1)}\}$$
and $\vec{D}\vec{D}$ is an absorbing state if
$$c < \min\{\frac{n_H}{2(N-1)}, \frac{n_L}{2(N-1)}\}\,.$$
If c does not satisfy one of these two conditions we can work out the following cycle by using the Δ-functions above:
$$\ldots \to \vec{C}\vec{D} \to \vec{D}\vec{D} \to \vec{D}\vec{C} \to \vec{C}\vec{C} \to \vec{C}\vec{D} \to \vec{D}\vec{D} \to \vec{D}\vec{C} \to \ldots$$
For c in an interval $[\min\{\frac{n_H}{2(N-1)}, \frac{n_L}{2(N-1)}\}, \max\{\frac{n_H+2n_L}{2(N-1)}, \frac{2n_H+n_L}{2(N-1)}\}]$ we will observe an absorbing set of strategy states. The imitation dynamics cycles between the four strategy states $\left\{\vec{C}\vec{D}, \vec{D}\vec{D}, \vec{D}\vec{C}, \vec{C}\vec{C}\right\}$, where a single agent plays C for two periods followed by D for two periods etc. The strategy states specific to the two type groups are shifted in time by one or three periods, depending on where in the cycle we start to count.

2.A.3 Decision Terms for Edge Players

The decision terms for an edge player of type H are

$$\begin{aligned}
\Delta^1_{H^E}(y_H, y_L) &= u_H(1, y_H, y_L) - u_L(0, y_H, y_L) \\
&= \frac{n_H + 2n_L + y_H - y_L}{2(N-1)} - c\,, && (2.16)\\
\Delta^2_{H^E}(y_H, y_L) &= u_H(1, y_H, y_L) - \frac{1}{2}[u_H(0, y_H, y_L) + u_L(0, y_H, y_L)] \\
&= \frac{n_H + 2n_L + \frac{1}{2}y_H - \frac{3}{2}y_L}{2(N-1)} - c\,, && (2.17)\\
\Delta^3_{H^E}(y_H, y_L) &= \frac{1}{2}[u_H(1, y_H, y_L) + u_L(1, y_H, y_L)] - u_H(0, y_H, y_L) \\
&= \frac{\frac{1}{2}n_H + \frac{3}{2}n_L + \frac{1}{2}y_H - \frac{3}{2}y_L}{2(N-1)} - c\,, && (2.18)\\
\Delta^4_{H^E}(y_H, y_L) &= u_L(1, y_H, y_L) - u_H(0, y_H, y_L) \\
&= \frac{n_L + y_H - y_L}{2(N-1)} - c\,. && (2.19)
\end{aligned}$$

For an edge player of type L we get

$$\begin{aligned}
\Delta^1_{L^E}(y_H, y_L) &= u_L(1, y_H, y_L) - u_H(0, y_H, y_L) \\
&= \frac{n_L + y_H - y_L}{2(N-1)} - c\,, && (2.20)\\
\Delta^2_{L^E}(y_H, y_L) &= u_L(1, y_H, y_L) - \frac{1}{2}[u_H(0, y_H, y_L) + u_L(0, y_H, y_L)] \\
&= \frac{n_L + \frac{3}{2}y_H - \frac{1}{2}y_L}{2(N-1)} - c, && (2.21)\\
\Delta^3_{L^E}(y_H, y_L) &= \frac{1}{2}[u_L(1, y_H, y_L) + u_H(1, y_H, y_L)] - u_L(0, y_H, y_L) \\
&= \frac{\frac{1}{2}n_H + \frac{3}{2}n_L + \frac{3}{2}y_H - \frac{1}{2}y_L}{2(N-1)} - c\,, && (2.22)\\
\Delta^4_{L^E}(y_H, y_L) &= u_H(1, y_H, y_L) - u_L(0, y_H, y_L) \\
&= \frac{n_H + 2n_L + y_H - y_L}{2(N-1)} - c\,. && (2.23)
\end{aligned}$$

2.A.4 Decision Terms for Double-Edge Players

For a double-edge player of type H the decision terms are

$$\Delta^1_{H^{EE}}(y_H, y_L) = \frac{n_L + y_H - y_L}{2(N-1)} - c\,, \tag{2.24}$$

$$\Delta^2_{H^{EE}}(y_H, y_L) = \frac{n_L + \frac{3}{2}y_H - \frac{1}{2}y_L}{2(N-1)} - c\,, \tag{2.25}$$

$$\Delta^3_{H^{EE}}(y_H, y_L) = \frac{\frac{1}{2}n_H + \frac{3}{2}n_L + \frac{3}{2}y_H - \frac{1}{2}y_L}{2(N-1)} - c\,, \tag{2.26}$$

$$\Delta^4_{H^{EE}}(y_H, y_L) = \frac{n_H + 2n_L + y_H - y_L}{2\,(N-1)} - c\,, \tag{2.27}$$

and for a double-edge player of type L they are

$$\Delta^1_{L^{EE}}(y_H, y_L) = \frac{n_H + 2n_L + y_H - y_L}{2(N-1)} - c\,, \tag{2.28}$$

$$\Delta^2_{L^{EE}}(y_H, y_L) = \frac{n_H + 2n_L + \frac{1}{2}y_H - \frac{3}{2}y_L}{2(N-1)} - c\,, \tag{2.29}$$

$$\Delta^3_{L^{EE}}(y_H, y_L) = \frac{\frac{1}{2}n_H + \frac{3}{2}n_L + \frac{1}{2}y_H - \frac{3}{2}y_L}{2(N-1)} - c\,, \tag{2.30}$$

$$\Delta^4_{L^{EE}}(y_H, y_L) = \frac{n_L + y_H - y_L}{2(N-1)} - c\,. \tag{2.31}$$

2.A.5 Proof of Proposition 5

We accomplish the proof of Proposition 5 in two parts. We first consider the absorbing states when an L-type introduces D.

L-type introduces D.

In the following we assume that the innovator is an interior player L^I. The absorbing states are the same when an L^{NE} or an L^E innovates D.

(a) If $c > \frac{3n}{2(N-1)}$, in the absorbing state all players play C because from (2.6) $\Delta_{L^I}(n, n-1) < 0$. Consequently, D dies out immediately.

(b) If $c \in \left(\frac{2n}{2(N-1)}, \frac{3n}{2(N-1)}\right)$ the absorbing state is $\vec{C}\vec{D}^*$. This interval can be divided into two subintervals. First, if $c \in \left(\frac{\frac{5}{2}n-1}{2(N-1)}, \frac{3n}{2(N-1)}\right)$ in the absorbing state all L^I play D, and the two L^E and all H-types play C. If $c < \frac{3n}{2(N-1)}$, interior players of type L imitate D because $\Delta_{L^I}(n, y_L) = \frac{3n}{2(N-1)} - c > 0$. (Note that $\Delta_{L^I}(n, y_L)$ does not depend on y_L). At one point in time, an L^E observes D. He adopts it if $\Delta^2_{L^E}(n, y_L) = \frac{\frac{5}{2}n - \frac{1}{2}y_L}{2(N-1)} - c > 0$. Since $\Delta^2_{L^E}$ depends negatively on y_L, $\Delta^2_{L^E}$ is increasing when the number of L^I using D increases. Thus, $\Delta^2_{L^E}(n, y_L)$ is maximal when $y_L = 2$. Consequently, if $\frac{\frac{5}{2}n-1}{2(N-1)} - c > 0$, the edge players of type L never adopt D, and an absorbing state is reached.

Second, if $c \in \left(\frac{2n}{2(N-1)}, \frac{\frac{5}{2}n-1}{2(N-1)}\right)$, in the absorbing state all H play C, all L^I play C, and the L^E cycle between D and C. If $c < \frac{\frac{5}{2}n-1}{2(N-1)}$, an L^E-type imitates D. The first H^E-agent observing D has the decision term $\Delta^4_{H^E}(y_H, y_L) = \frac{n+y_H-y_L}{2(N-1)} - c$. He adopts D if $\Delta^4_{H^E}(y_H, y_L) > 0$. Since all $H-$types still use C we have $y_H = n$. Let us first assume that c is such that he does not imitate D. Then, the number L-types adopting D is increasing because their decision term does not depend on y_L. Consequently, y_L decreases to 0 implying $\Delta^4_{H^E}(n, 0) = \frac{N}{2(N-1)} - c$.[12] Thus, if $\frac{2n}{2(N-1)} < c < \frac{\frac{5}{2}n-1}{2(N-1)}$ all H-types play C and the L^E-types cycle between C and D because $\Delta^1_{L^E}(n, 0) < 0$ and $\Delta^2_{L^E}(n, 2) > 0$.

(c) If $c < \frac{2n}{2(N-1)}$, in the absorbing state all players play D. If $c < \frac{2n}{2(N-1)}$, $\Delta^4_{H^E}(y_H, y_L) > 0$. Then, from (2.10) all other H-types will also imitate D. Note that L-types continue to use D because they do not observe strategy C in their information set anymore. Consequently, in the absorbing state all players play D.

H-type introduces D.

(i)

If an H^I introduces D, we have to consider the intervals $c \in \left(0, \frac{n}{2(N-1)}\right)$

[12] Note that y_L can only become 1 and not 0 if n is even. Start counting outward from an L^I inventing D. Since n is even, it takes an uneven number of periods from the period the the first L^E observes D until the second L^E observes D. As explained in the text, the L^E cycle between C and D for the interval under consideration. Hence the two L^E do not play the same strategies in any given period. This shifts the lower bound of the respective interval to $c \in \left(\frac{2n-1}{2(N-1)}, \frac{\frac{5}{2}n-1}{2(N-1)}\right)$ for even n. In the Proposition we state the result for odd n.

and $c \in \left(\frac{n}{2(N-1)}, 1\right)$.

(a) If $c > \frac{n}{2(N-1)}$, then $\Delta_{H^I}(n-1,n) < 0$ and D is extinguished in the second period. The absorbing state is $\vec{C}\vec{C}$.

(b) If $c < \frac{n}{2(N-1)}$, then $\Delta_{H^I}(n-1,n) > 0$, and the H^I that observe D will play it in the second period. In the following periods, more and more H^I will switch to D since Δ_{H^I} is unchanged as long as only H-types observe D. At some point an H^E will observe D. His decision term is $\Delta^2_{H^E}$, which from (2.10), is always higher than $\Delta_{H^I}(n-1,n)$. So he will adopt. The next player to consider is L^E. His decision term is $\Delta^4_{L^E}(y_H, n)$. He also adopts D because $\Delta^4_{L^E}(y_H, y_L) > \Delta_{H^I}(y_H, y_L) > 0$. Now all L^I adopt D too, since $\Delta_{L^I} > 0$ for $c < \frac{n}{n(N-1)}$. The last question we have to answer is whether the decreasing y_H or y_L can stop the spreading in either of the two groups. For the H-types this is not the case because Δ_{H^I} is decreasing in y_L. For the L-types we find that if $c < \frac{n}{2(N-1)}$, then Δ_{L^I} is always positive. Consequently, if $c < \frac{n}{2(N-1)}$, the absorbing state is $\vec{D}\vec{D}$.

(ii)
If an H^N invents D, one H^I and one H^E observe D. For this reason, it is convenient to divide the proof into three subsections: we consider the cost intervals $\left(0, \frac{n}{2(N-1)}\right)$, $\left(\frac{n}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$, and $\left(\frac{2n-\frac{1}{2}}{2(N-1)}, 1\right)$ separately.

(a) If $c < \frac{n}{2(N-1)}$, one can show that all decision terms are strictly positive. Consequently, the absorbing state is $\vec{D}\vec{D}$.

(b) For $\frac{n}{2(N-1)} < c < \frac{2n-\frac{1}{2}}{2(N-1)}$, we find several absorbing states. We derive them one by one. In the first period we have $\Delta^2_{H^E}(n-1,n) > 0$ but $\Delta_{H^I}(n-1,n) < 0$. In period $t=2$, H^E is thus the only agent playing D, since H^N who invented D abandons it (he decides according to Δ_{H^I}) and H^I does not adopt it. So now L^E, H^E and H^N face the decision to imitate D or not. Their decision terms are $\Delta^4_{L^E}$, $\Delta^2_{H^E}$ and Δ_{H^I}, valued at $(n-1,n)$. With $c \in \left(\frac{n}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$ we have

$$\begin{aligned}
\Delta^4_{L^E}(n-1,n) &= \frac{3n-1}{2(N-1)} - c > 0 \\
\Delta^2_{H^E}(n-1,n) &= \frac{2n-\frac{1}{2}}{2(N-1)} - c > 0 \\
\Delta_{H^I}(n-1,n) &= \frac{n}{2(N-1)} - c < 0\,.
\end{aligned}$$

So we know that in period $t=3$ L^E as well as H^E play D and L^I, L^E, H^E, and H^I have to decide to adopt or abandon D. By looking at these four decision terms (Δ_{L^I}, $\Delta^3_{L^E}$, $\Delta^3_{H^E}$, and Δ_{H^I}) evaluated at $(n-1,n-1)$, it becomes clear that the process now is not unique for the whole interval of $c \in \left(\frac{n}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$.

$$\begin{aligned}
\Delta_{L^I}(n-1,n-1) &= \frac{3n-2}{2(N-1)} - c > 0 \\
\Delta^3_{L^E}(n-1,n-1) &= \frac{3n-1}{2(N-1)} - c > 0 \\
\Delta^3_{H^E}(n-1,n-1) &= \frac{n+1}{2(N-1)} - c \gtrless 0 \\
\Delta_{H^I}(n-1,n-1) &= \frac{n+2}{2(N-1)} - c \gtrless 0\,.
\end{aligned}$$

The H^E who has the choice will only adopt D if $c < \frac{n+1}{2(N-1)}$. In this case, D will be played by the whole population: all H^I and L^I have a positive decision term such that y_H and y_L decline by one each period. Also the last two agents confronted with D, H^E and L^E imitate D, because

$$\begin{aligned}
\Delta^2_{H^E}(1,1) &= \frac{3n-1}{2(N-1)} - c > 0 \\
\Delta^2_{L^E}(1,1) &= \frac{n+1}{2(N-1)} - c > 0\,.
\end{aligned} \tag{2.32}$$

So for $\frac{n}{2(N-1)} < c < \frac{n+1}{2(N-1)}$, the absorbing state is $\vec{D}\vec{D}$.

It is clear that for a slightly higher c, $c \in \left(\frac{n+1}{2(N-1)}, \frac{n+2}{2(N-1)}\right)$, D spreads too in both populations, but once all agents but two play D, L^E will not adopt it, according to (2.32). Because of

$$\Delta_{L^I}(0,1) = \frac{n}{2(N-1)} - c < 0\,,$$

C will be imitated by the L^{NE} next to L^E in the following period, in which case L^E switches to D (the decision term for L^E now is $\Delta^3_{L^E}$, which is positive for the c under consideration). All L-players but the edge players adopt C then. The edge players' decision term $\Delta^2_{L^E}$ cannot be negative for the c under consideration and thus the H-types are cut off the strategy C. For the interval $\frac{n+1}{2(N-1)} < c < \frac{n+2}{2(N-1)}$ we thus find the absorbing state $\vec{D}\vec{C}^*$.

Rising c again, $c \in \left(\frac{n+2}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$, in period $t = 4$ none of the H-players will imitate D and it will spread in group L. Once y_L is small enough, group H will adopt strategy D. The absorbing state here is $\vec{D}\vec{D}$.

(c) If $\left(\frac{2n-\frac{1}{2}}{2(N-1)}, 1\right)$, we have $\Delta^2_{H^E}(n-1, n) < 0$ and $\Delta_{H^I}(n-1, n)$ respectively. So all players give up D in the second period and the absorbing state is $\vec{C}\vec{C}$.

(iii)
If an H^E introduces D, an L^E and an H^I and the introducing H^E can decide to play D or not in the second period. The respective decision terms are

$$\begin{aligned}\Delta^4_{L^E}(n-1,n) &= \frac{3n-1}{2(N-1)}\\ \Delta^2_{H^E}(n-1,n) &= \frac{2n-\frac{1}{2}}{2(N-1)}\\ \Delta_{H^I}(n-1,n) &= \frac{n}{2(N-1)}.\end{aligned}$$

It is convenient to look at the cost intervals $\left(0, \frac{n}{2(N-1)}\right)$, $\left(\frac{n}{2(N-1)}, \frac{2n-\frac{1}{2}}{2(N-1)}\right)$, $\left(\frac{2n-\frac{1}{2}}{2(N-1)}, \frac{3n-1}{2(N-1)}\right)$, and $\left(\frac{3n-1}{2(N-1)}, 1\right)$ separately.

(a) If $0 < c < \frac{n}{2(N-1)}$, the absorbing state is $\vec{C}\vec{C}$ because all decision terms are positive.

(b) If $\frac{n}{2(N-1)} < c < \frac{2n-\frac{1}{2}}{2(N-1)}$ then $\Delta^2_{H^E}(n-1,n) > 0$ and $\Delta^4_{L^E}(n-1,n) > 0$, and $\Delta_{H^I}(n-1,n) < 0$. So in the second period H^E and L^E only play D. This situation is what we have analyzed in Part (ii), subsection (b) of this proof. We can therefore take over these results and conclude the similar dynamics for this range of the costs.

(c) If $\frac{2n-\frac{1}{2}}{2(N-1)} < c < \frac{3n-1}{2(N-1)}$) only L^E plays D in the second period. All L^I adopt D in the following because Δ_{L^I} is always negative for the given cost interval. The spread of strategy D among the L-types will not affect the H-types strategy choice, they keep playing C because $\Delta^4_{H^E} < 0$ for the cost interval under consideration. As discussed before, the L^E either cycle between C and D or play D, the absorbing state for this cost interval is $\vec{C}\vec{D}^*$.

(d) If $\frac{3n-1}{2(N-1)} < c < 1$, the strategy D becomes extinct and the absorbing state is $\vec{C}\vec{C}$, because all decision terms are negative.

2.A.6 Proof of Proposition 6

We first assume that an H-type plays D in period $t = 1$. If $\Delta^4_{H^{EE}} > 0$ he will keep playing D and his two neighbors adopt D if $\Delta^2_{L^{EE}} > 0$. We have $\Delta^4_{H^{EE}}(n-1,n) > \Delta^2_{L^{EE}}(n-1,n)$ and so D becomes extinct if $c > \frac{3n-1}{2(N-1)}$. For $\frac{2n-\frac{1}{2}}{2(N-1)} < c < \frac{3n-1}{2(N-1)}$ the equilibrium is the initial strategy distribution, because the neighbors of the inventor do not adopt D, but the H^{EE} keeps playing it.

We focus on $c < \frac{2n-\frac{1}{2}}{2(N-1)}$ now. As just seen, in $t = 2$ only the innovator H^{EE} and his two neighbors play D. Four players can choose between C and D now, two H^{EE} (playing C) that have the decision term $\Delta^2_{H^{EE}}$ and two L^{EE} (playing D) that have the decision term $\Delta^3_{L^{EE}}$. Since $\Delta^2_{H^{EE}}(n-1,n-2) > \Delta^3_{L^{EE}}(n-1,n-2)$ and $\Delta^2_{H^{EE}}(n-1,n-2) > 0$, there are two cases to be distinguished $t = 3$: either all of them adopt D, or the two H^{EE} adopt D and the two L^{EE} do not. The two cases will actually lead to the same absorbing state: the two L^{EE} will have the decision term $\Delta^1_{L^{EE}}(y_H, y_L)$ in period $t = 4$ which is greater than $\Delta^2_{H^{EE}}(y_H, y_L)$ for all y_H and y_L, so they will adopt D then. Thus, this second case does not influence the absorbing state. It is clear that if the two H^{EE} choose to play D, the L^{EE} will do the same one period later. We neglect it from now on.

There are only two different kind of strategy states in this dynamics. The first one is of the kind of period $t = 3$: two H^{EE} with decision term $\Delta^2_{H^{EE}}$ and two L^{EE} with $\Delta^3_{L^{EE}}$ choose between D and C. We call this state L-dominated because more L-types than H-types play D (to be precise: $y_L - y_H = 1$). Picture period $t = 3$ to get the intuition. In $t = 3$, three agents play D, these are located as follows:

$$\underbrace{...HLHLH}_{C}\underbrace{LHL}_{D}\underbrace{HLHLHL...}_{C}.$$

The second kind of strategy state is of the kind of period $t = 4$: two L^{EE} with decision term $\Delta^2_{L^{EE}}$ and two H^{EE} with decision term $\Delta^3_{H^{EE}}$ have to decide between the two strategies. We call this state H-dominated because more H-types than L-types play D ($y_H - y_L = 1$).

$$\underbrace{...HLHL}_{C}\underbrace{HLHLH}_{D}\underbrace{LHLHL...}_{C}.$$

Note that for any positive integer x, $\Delta^2_{L^{EE}}(x-2, x-1)$ as well as $\Delta^3_{H^{EE}}(x-2, x-1)$ are larger than $\Delta^2_{H^{EE}}(x, x-1)$. So if we have reached an H-dominated state, we will also reach the L-dominated state that features two more L-types playing D.

We conclude that the absorbing state is reached when $\Delta^2_{H^{EE}}$ is negative ($\Delta^2_{H^{EE}}$ depends positively on y_H). So only an L-dominated state can be an absorbing state. The interpretation is that it will always be an H-type that stops the spread of D. An H-type will at one point halt the spread of D and act as a blocker to the L-types playing C who would adopt D if they would observe D. The higher c is, the earlier the spread of D is halted, that means, the less agents use D in the absorbing state. We conclude that there exist many different absorbing states, depending on population size N and the costs c.

We calculate the number of agents playing C in the absorbing state. Since the absorbing state is L-dominated, we can substitute y_L with $y_H - 1$.

$$\begin{aligned} \Delta^2_{H^{EE}} &= \frac{n + y_H + \frac{1}{2}}{2(N-1)} - c < 0 \\ &\rightarrow \; y_H < 2c(N-1) - n - \frac{1}{2}. \end{aligned}$$

The solution y_H to this inequality must be an odd (even) number if n is even (odd). This comes from the fact that whenever H^{EE} have the decision term $\Delta^2_{H^{EE}}$, there is an odd number of H-types playing D. This is because y_H decreases by steps of two.

The same arguments apply if the innovator is an L-type.

Chapter 3

Informational Deficits and Strategy Clustering

3.1 Introduction

In Chapter 2 we have studied the role of information and the role of heterogeneity for a competition in which agents could either adopt a fair or an illegal strategy. We did so by means of an evolutionary game with imitation dynamics. We briefly repeat its basic structure. The underlying game represents the competition of two agents for a prize of value w. The agents choose between two strategies, they can either act legally (playing fair) or illegally (cheating). An agent acting illegally wins the prize when playing against a fair playing agent, but bears the costs c. An agent's payoff in every period is the sum of outcomes received from playing the game against all other agents. We have presented a model variant with homogenous agents and one with heterogenous agents. Heterogeneity implies that agents either are of high or of low type. The high types have a natural advantage over the low types in that they win the prize when competing against a low type with the same strategy. However, a low type wins the prize against a high type if the low type cheats and the high type plays fair. The agents are located on a circle which allows us to define local and global information. These two concepts constitute different information sets for the agents. We assume that agents imitate the observed strategy with the highest average payoff. For this setup, we have formally described the absorbing states of the imitation dynamics for initial situations where only one agent has the illegal activity at his disposal. Furthermore, we have explained how the type of the innovator, the segregation of the population, and the size of information sets influence the absorbing states.

We have argued, that the model variant with heterogenous agents and local information depicts the information deficits – scarcity and non-verifiability – prevailing for decisions concerning the adoption of illegal activities. In this chapter we take up this model variant in order to describe its absorbing states for a broader set of initial strategy states. We again interpret the results as the outcomes from imitation dynamics of illegal activities under information deficits.

Instead of calculating the imitation dynamics explicitly like we have done in the last chapter, we here display the characteristics of absorbing strategy states found by numerical simulations. The initial strategy states used for the simulations are drawn at random. So in contrast to the last chapter, we do not restrict ourselves to initial strategy states with only one agent acting illegally. The model therefore applies to situations where illegal activities emerge at diverse locations in a population simultaneously. Or similarly to situations where a population newly forms up both from agents that have knowledge of the illegal activity and from agents that do not. One application of the setup treated here is corruption in a government department. People are hired from different work environments and may or may not have been in contact with corruption in their precedent occupation. It is easily conceivable that those that have been involved in corrupt manipulation before will bring in this behavioral strategy to the newly formed group. If they then build up confidence relationships (corresponding to local information) independently from their knowledge about corruption, our model applies.

We also abandon the assumption on specially ordered type distributions, that is, we allow for more type distributions than just maximally and minimally segregated populations. As well as the initial strategy states, we now draw the type distribution of a population randomly.

We are interested in answering the following questions concerning the characteristics of absorbing states: Do high types or low types use the illegal strategy more often at a given cost level? Do neighbors tend to play the same strategies regardless of their types, i.e. are strategies played in clusters? If there is a clustering of strategies, where do we have to expect these clusters to be with respect to the type distribution? Does the initial share of agents

playing the illegal activity matter? What kind of type distributions exhibit a low share of agents playing the illegal activity in the absorbing states? The simulation results are displayed graphically, and in dependence of the cost parameter c. In most of the figures, the dependent variable is the expected frequency of the legal strategy in the absorbing states.

The chapter is organized as follows. In Section 3.2 we briefly explain the applied simulation method. Additionally we explain the standard settings for our simulations. In section 3.3 we describe the characteristics of the absorbing strategy states when the type distributions are drawn randomly. First, we investigate how the probability that a high type acts illegally differs from the probability that a low types acts illegally. Second, we analyze what kind of positions in the type distribution account for the last result. Third, we show that the two strategies are played in clusters in our model. In section 3.4, we depart from the standard setting that have been used for the simulations so far. We vary the share of agents playing fair initially, study initial strategy states with clusters, and look at populations where high types outnumber low types and vice versa. In Appendix 3.A we exemplarily display the imitation dynamics of our evolutionary game for a randomly drawn initial strategy state and a randomly drawn type distribution. We list the played strategy state for every period such that the reader gets an idea of how the imitation dynamics changes the initial strategy state to the absorbing state.

3.2 Monte Carlo Simulations

We apply a Monte Carlo Method (see e.g. Greene, 2000; Judd, 1998) to study the characteristics of the absorbing states.

3.2.1 The Procedure

Assume that we are interested in a characteristic γ of the absorbing states.[1] For our model, the mean of γ, $E[\gamma]$, is the most important moment to generally describe the characteristic. Since we allow for random initial strategy

[1]Note that *absorbing state* always refers to the strategy state, and not the type distribution. The type distribution does not change in the evolutionary game under consideration.

states, we can view a characteristic γ_r of an absorbing state arising from a given initial strategy state as a realization of a stochastic process. Producing a large number of random realizations allows us to calculate the average

$$\overline{\gamma}_R = \frac{1}{R}\sum_{r=1}^{R}\gamma_r$$

of the characteristics γ_r, $r = 1, ..., R$. According to a strong law of large numbers, this average $\overline{\gamma}_R$ converges to $E[\gamma]$, the mean of γ (Amemiya, 1998). The procedure consists of five steps.

1. We draw the initial strategy state and the type distribution randomly. Alternatively we might determine some properties of either of the two, depending on what kind of characteristics of the absorbing states we aim to analyze. Note that the type distribution remains unchanged for the computation of an absorbing state.

2. We compute the strategy state of the next period. This is done by counting through the type distribution and the initial strategy state which yields n_H, n_L, y_H, and y_L. These numbers are used to evaluate the decision term (2.2) of every agent, which defines his strategy choice of the next period according to the imitation rule (2.3). The strategy state of the next period is defined herewith.

3. The second step is repeated until the strategy state does not change anymore. If this is the case, we have found an absorbing state. Alternatively, we stop repeating the second step, if we find a strategy state that has been played in an earlier period already. Such a state is an element of a set of absorbing states.

4. The absorbing strategy state (or the set of absorbing states) is analyzed on the characteristic γ under consideration (e.g. size of strategy clusters). The characteristics is listed as the number γ_r.

5. All above steps are repeated R times.

6. We average over the results collected and get the average of the characteristics, $\overline{\gamma}_R$.

Since we draw randomly, our averages converge to the means almost surely. Note that the above procedure needs to be completed for every value of c.

3.2.2 Standard Simulation Settings

In this section we display the standard settings for our simulations. Unless otherwise noted, the following settings are used for the simulations. In order to keep the descriptions as simple as possible, we adhere to the notation of Chapter 2 in the following.

We accomplish our simulations with a population size of 40 players, i.e. $N = 40$.

We draw the type distribution simultaneously with an initial strategy state. For the type distribution we assume, that the probability that an agent is an H-type is $\frac{1}{2}$. This implies that

$$E[n_H] = E[n_L] = \frac{N}{2}\,.$$

We deviate from this assumption in section 3.4.2.

For the initial strategy state we assume that the probability that an agent plays D is the same as that he plays C. This implies that

$$E[y_H] = E[n_H - y_H] \quad \text{and} \quad E[y_L] = E[n_L - y_L]\,.$$

We analyze deviations from this assumption in section 3.4.1.

For the analysis of a specific characteristic γ, we generate at least 500 random realizations γ_R, i.e. $R \geq 500$. That means, that we draw at least 500 pairs of type distributions and initial strategy states, wherefrom we calculate the absorbing states.

We perform all simulations for 101 values of c, $c = 0, 0.01, 0.02, ..., 0.99, 1$.

3.3 Characterizing the Absorbing States

In this section we display the results from simulations that are run with the standard settings described in section 3.2.2. We focus on three characteristics of the absorbing states: First, we describe if either H-types or L-types play D more frequently at given costs c. Second, we define the term *position* in the type distribution and explain, how the agents on different positions vary in their probability to play C in the absorbing state. Third, we define *strategy clusters* and compare the number of strategy clusters in an absorbing state with the number of strategy clusters in a random strategy state.

Note that the expected share of agents playing a given strategy in the absorbing state can always be interpreted as the probability of a single agent to play this strategy. We switch between these two interpretations without mentioning their analogy in the following.

3.3.1 Discerning Strategy Choices of H- and L-Types

Before looking at the strategy choices for H-types and L-types separately, we want to know how many agents of the whole population choose D for given costs c. The (average) share of C-playing agents for the costs c, $\sigma(c)$, is defined as

$$\sigma(c) = \frac{1}{R} \sum_{r=1}^{R} \frac{y_H^r(c) + y_L^r(c)}{N}.$$

In Figure 3.1 $\sigma(c)$ is displayed as the black line. We summarize our observations concerning $\sigma(c)$ from Figure 3.1 in Result 1.

Result 1 *For randomly drawn type distributions and randomly drawn initial strategy states the expected share of agents playing C increases smoothly and monotonously.*

The shares of H-types and L-types choosing C differ for almost all costs c.

$$\sigma_H(c) = \frac{1}{R} \sum_{r=1}^{R} \frac{y_H^r(c)}{n_H^r} \quad \neq \quad \sigma_L(c) = \frac{1}{R} \sum_{r=1}^{R} \frac{y_L^r(c)}{n_L^r}$$

In Figure 3.1 $\sigma_H(c)$ (red line) and $\sigma_L(c)$ (blue line) are displayed. We summarize our insights from Figure 3.1 in the next result.

Result 2 *There exists a cost-threshold $\tilde{c}$, where $\sigma_H(\tilde{c}) = \sigma_L(\tilde{c})$. For $c < \tilde{c}$, the share of H-types playing C is smaller than the share of L-types playing C, i.e. $\sigma_H(c) < \sigma_L(c)$. For $c > \tilde{c}$, the opposite is true.*

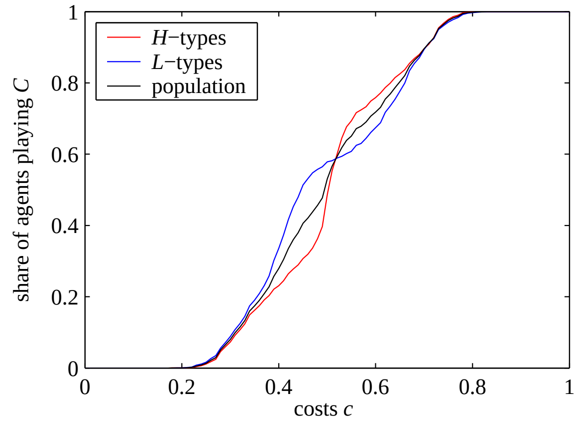


Figure 3.1: The share of agents playing C.

Of course we have that $\sigma_H(c) = \sigma_L(c) = 0$ $(= 1)$ for very low (high) costs c. For the range of c where both strategies are played, H-types exhibit a greater probability to play D than L-types if costs c are low. For high costs however, L-types play D with the greater probability.

We have described analytically the absorbing state of maximally and minimally segregated populations in Chapter 2. Maximally and minimally segregated populations are the two extreme cases of type distributions. However, Result 2 is quite different from what we have found in Propositions 5 and 6. In a maximally segregated population, the shares of H-types and L-types playing C have only been different for a small cost interval (see Figure 2.3). In a minimally segregated population, the shares of agents playing C has even been the same for both types at all costs c. In contrast, we now observe that H- and L-types exhibit unequal shares of agents playing C. This motivates us to analyze, which positions in the type distributions account for this result.

3.3.2 Different Positions in the Type Distribution

In Chapter 2 we have analyzed the decision terms of different positions in the type distribution: we have distinguished interior players, edge players, and double-edge players. We refine this idea of positions by attaching a position number g to every agent. If an agent has position number g, then there are exactly g other agents between him and the next agent of different type. Consequently, an edge player is assigned the position number 0. An agent whose neighbor is an edge player but is not an edge player himself gets the position number 1. An example clarifies our intention. In the type distribution

$$HHHLHHLLHHHHHL\,,$$

the players are assigned the position numbers

$$\underbrace{H}_{0}\underbrace{H}_{1}\underbrace{H}_{0}\underbrace{L}_{0}\underbrace{H}_{0}\underbrace{H}_{0}\underbrace{L}_{0}\underbrace{L}_{0}\underbrace{H}_{0}\underbrace{H}_{1}\underbrace{H}_{2}\underbrace{H}_{1}\underbrace{H}_{0}\underbrace{L}_{0}\,.$$

In the following we call a player with position number g a g-player. We abbreviate the number of H-types (L-types) with position number g by n_{H^g} (n_{L^g}). Analogously, y_{H^g} (y_{L^g}) are the H-types (L-types) with position number g, that play C. We simulate our model again and list the share of agents playing C for every position number separately.

$$\sigma_H^g(c) = \frac{1}{R}\sum_{r=1}^{R}\frac{y_{H^g}^r(c)}{n_{H^g}^r} \quad \text{and} \quad \sigma_L^g(c) = \frac{1}{R}\sum_{r=1}^{R}\frac{y_{L^g}^r(c)}{n_{L^g}^r} \quad \forall\, g\,.$$

The population size is $N = 60$ for these simulations. In Figure 3.2 we plot $\sigma_{H^g}(c)$ and $\sigma_{H^g}(c)$ for $g = 0, ..., 5$.[2] We summarize our observations in the next result.

Result 3 *H-types as well as L-types with position number 0 (edge players and double-edge players) exhibit probabilities to play C that are different from the players with other position numbers.*[3] *Players of one type with position numbers greater than 0 exhibit similar probabilities to play C.*

We discuss the H-types first. The probability that a 0-player plays C rises more or less linearly with the costs c for the interval $c \in [0.2, 0.8]$. This is not true for agents with a different position number. The higher the position number of a player is (the further away an H-type is from the

[2] In the hundred thousands of type distributions we drawed randomly with $N = 60$, we did not observe any agent with position number 6.

[3] Statistically significant for almost all $c \in [0, 1]$.

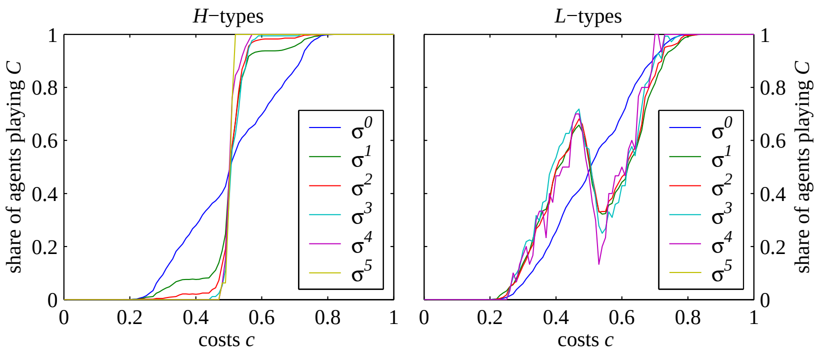


Figure 3.2: Share of agents playing C for different positions.

next L-player), the more extreme his behavior becomes. An H-type with position number 5 does practically not react to c except that his strategy choice tips over at a $c = 0.5$. As long as c is below that threshold, the probability that he plays D is almost 1. For costs above this threshold he most likely plays C.

The pattern is different for the L-types. While the 0-players again react very smoothly to c, the behavior of the agents with other positions numbers is a little bit more complicated. They tend to play C more often than the 0-players for $c < \frac{1}{2}$, and more rarely than the edge-players for high $c > \frac{1}{2}$. Their probability to adopt C increases quite sharply for low costs, but drops drastically for costs around $\frac{1}{2}$. Note that there are not enough players with position numbers greater than 0 in the population, such that the population behavior would mirror this non-monotonicity in the probability to play C (compare the blue line in Figure 3.1).

We plot σ_H^0 and σ_L^0 in Figure 3.3. Are the probabilities for 0-players to play C significantly different for the two types? They actually are, except from a small number of cost levels. It has to be noted though, that σ_H^0 and σ_L^0 cross three times (see Figure 3.3) for moderate c which does not comply with the behavior of other positions at all. And in comparison with the other agents, they exhibit very similar probabilities of playing C.

We conclude that those agents who exchange information with agents of the other type (0-positions), disclose strategy choices that differ considerably from those only exchanging information with their own kind (all other

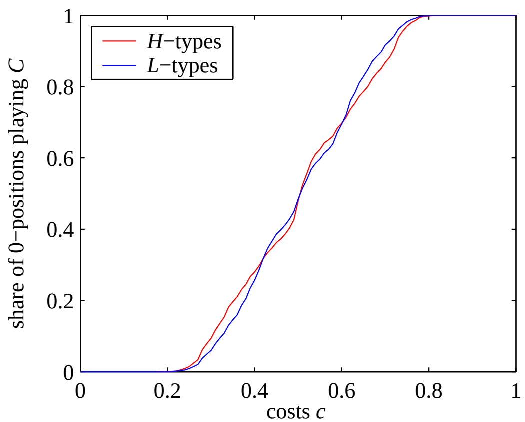


Figure 3.3: Shares of 0-positions playing C.

positions). This is true for all non-negligible cost intervals. Although the 0-players of the two types feature statistically different probabilities to play C, the two probabilities run closely. Obviously, for 0-position, the type looses its (otherwise crucial) impact on the expected probability to choose a strategy.

3.3.3 Strategy Clusters

In the previous section, we deal with the probabilities agents with different positions in the type distribution exhibit in playing C. We now neglect the type distribution and address a different characteristic of the absorbing states. We analyze if the absorbing states are similar to random strategy states or if they exhibit a certain pattern. For this reason, we define a *strategy cluster*. Remember that the players are located on a circle.

Definition 1 *A cluster is a sequence of players choosing the same strategy in a strategy state.*

An example clarifies the concept. We look at the following strategy state

$$DCDCDDDDCDDDCDCCDDCCCCC\,.$$

We divide the strategy state into the minimum number of parts that only contain one strategy (these are the clusters). Then we count through, which

yields the number of clusters in a strategy state.

$$\overbrace{D}^{1}\underbrace{C}_{2}\overbrace{D}^{3}\underbrace{C}_{4}\overbrace{DDDD}^{5}\underbrace{C}_{6}\overbrace{DDD}^{7}\underbrace{C}_{8}\overbrace{D}^{9}\underbrace{CC}_{10}\overbrace{DD}^{11}\underbrace{CCCCC}_{12}$$

It is clear that the number of clusters always is an even number. Additionally, for a population with N players, the number of clusters and the size of the clusters are negatively related. We simulate our model as before and list the number of clusters for every absorbing state. From these, we calculate the average number of clusters for the costs c, $\overline{\eta}_R(c)$. Again, $\overline{\eta}_R(c)$ converges to $\eta(c)$. The average numbers of clusters in the model are compared with two other series. The first concept we compare $\eta(c)$ with, is the *potential number of clusters* for costs c. The second concept we use for the comparison, is the *random number of clusters* for costs c.

The *potential number of clusters* for costs c, $\eta^{pot}(c)$, denotes the number of clusters that could maximally appear at costs c in our model. In order to calculate $\eta^{pot}(c)$, we have to take the average number of agents playing each strategy at the costs c into account.[4] The maximal number of clusters is

$$\eta^{pot}(c) = 2 \cdot \min\{y_H(c) + y_L(c)\,,\, N - y_H(c) - y_L(c)\} \quad \forall c\,.$$

The reason is, that always the strategy played by fewer agents gives the upper bound for the number of clusters possible.

The *random numbers of clusters*, $\eta^{ran}(c)$, is the average number of clusters in strategy states drawn randomly. Analogously to the case of $\eta^{pot}(c)$, we use the share of agents playing C-players given by the model for drawing the random strategy states. For all costs c, we draw R random strategy states, where the probability of an agent to play C is equal to

$$\frac{y_H(c) + y_L(c)}{N}\,.$$

[4]We calculate the average number of C-playing agents for all cost levels in our model:

$$y_H(c) = \frac{1}{R}\sum_{r=1}^{R} y_H^r(c) \quad \text{and} \quad y_L(c) = \frac{1}{R}\sum_{r=1}^{R} y_L^r(c)\,.$$

The average number of C-playing agents is therefore $y_H(c) + y_L(c)$ for costs c.

We then count the clusters in these random strategy states and calculate their average, which gives us $\eta^{ran}(c)$.

In Figure 3.4 we display the results. We notice that the absorbing states

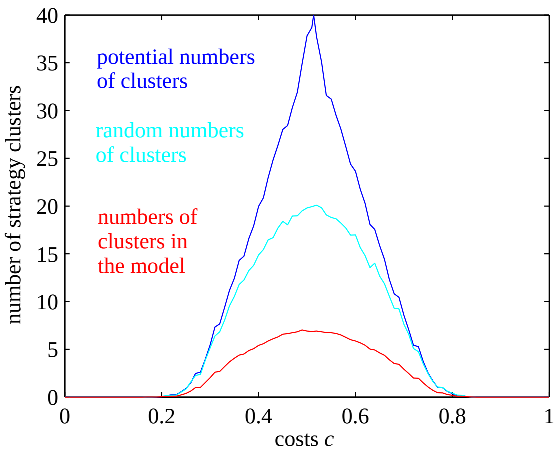


Figure 3.4: Comparing $\eta(c)$ with $\eta^{pot}(c)$ and $\eta^{ran}(c)$.

of our model are characterized by only a fraction of the clusters that could maximally appear. As well, the absorbing states show only a fraction of the clusters randomly drawn strategy states exhibit. Since we look at a population of a fixed number of players, we can conclude that at least some of the clusters must be relatively large in the absorbing states. We conclude the observations from Figure 3.4 in Result 4.

Result 4 *The absorbing strategy states of our model exhibit only a fraction of clusters a randomly drawn strategy state does. They are therefore characterized by relatively large strategy clusters.*

In Figure 3.4 we observe, that the numbers of clusters are symmetric with respect to c. So for the formation of clusters it does not matter, what strategy is played more frequently in the absorbing state. Neither it is important, if H-types or L-types constitute the majority of the agents playing C. Taking into account that the type distribution is drawn at random, this supports the interpretation that there exists a non-negligible number of agents for which the strategies played by neighbors are more important

than their actual type for the strategy choices. The probability that an agent has a neighbor using the same strategy as he does, is much larger in our model than it is if we draw their strategies randomly. This is true though the type distribution is drawn at random. We can therefore say that strategy choices are contagious in our model.

The reason for the relatively big strategy clusters in the absorbing states is, that 0-players (edge and double-edge players) over- and underestimate the value of a strategies in certain strategy states. The degree of how this feature determines the absorbing states is quite large: The model's absorbing states feature only 35% to 45% of the numbers of clusters of an equivalent random strategy state.

3.4 Varying the Standard Settings

In this section we vary the standard settings presented in section 3.2.2. First, we consider initial strategy states with uneven shares of the two strategies. This allows us to analyze, how an increasing share of agents playing C affects the frequency of C in the absorbing states. Additionally, we describe how strategy clusters in the initial strategy states influence the frequency of C in the absorbing states. Second, we vary the type distribution and describe how this impacts the frequency of strategy C in the absorbing states.

The considerations in this section could also be viewed as the analysis of policy implications. The requirement is that a policy maker can either control the initial strategy state or the type distribution. In this sense, our investigations provide information how to choose initial strategy states or type distribution in order to targets a low frequency of the illegal strategy in the absorbing states.

3.4.1 Varying the Initial Strategy State

We first investigate how the characteristics of an initial strategy state affect the frequency of C in the absorbing state. We do not make any assumptions on the type distribution and draw it randomly as described in section 3.2.2. This means that an agent's probability to be an H-type or an L-type is $\frac{1}{2}$ in both cases. We address the following two questions: Does the share of agents playing C in an initial strategy state significantly influence the frequency of C in the absorbing state? Do strategy clusters in the initial strategy state have an effect on the frequency of C in the absorbing state?

In order to answer the first question, we draw initial strategy states with different frequencies of C. We denote the frequency of strategy C in the initial strategy state by σ_{ISS}. For our simulations, we draw R initial strategy states with the property

$$E[\sigma_{ISS}] = a \quad \text{with} \quad a \in [0,1]$$

where we attribute a the values $a = 0.1, 0.2, ..., 0.9$. From these initial strategy states, we compute the absorbing states and evaluate their frequency of C, which we abbreviate by $\sigma^a(c)$. We display the simulation results in Figure 3.5. We depict our conclusion from Figure 3.5 in Result 5.

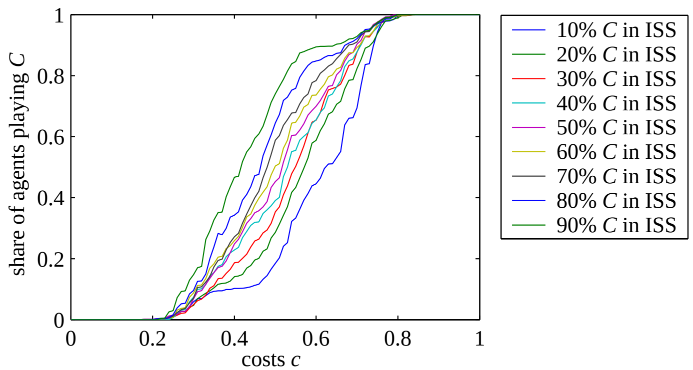


Figure 3.5: Different frequencies of C in initial strategy states.

Result 5 *The higher the share of agents playing C in the initial strategy state, the higher is the expected frequency of C in the absorbing state.*

However, the differences of initial strategy states with respect to their frequencies of C, are not mirrored to the same extend in their absorbing states. Let us choose $\sigma^{0.1}(c)$ and $\sigma^{0.9}(c)$ as an example. While the according initial strategy states differ by 80% of agents playing C, we do not observe a difference bigger than 57% of $\sigma^{0.9}(c) - \sigma^{0.1}(c)$ for any c.

Nevertheless, a policy maker targeting a low frequency of D in the absorbing state would aim to implement initial strategy states with the lowest σ_{ISS} possible.

We turn to the second question raised above. In the following we aim to describe how strategy clusters in the initial strategy states effect the frequency of C in the absorbing states.

We perform our analysis with initial strategy states characterized by $E[\sigma_{ISS}] = \frac{1}{2}$, that is, initial strategy states exhibit the same expected number of C- and D-players. Again, the type distribution is drawn randomly, with the expected ratio $E[\frac{n_H}{n_L}] = \frac{1}{2}$. For our simulations, we assign a certain number of agents one of the two strategies directly and only draw the strategies of the remaining agents randomly. By doing so, we assure that the initial strategy states contain a strategy cluster of a given size. The size of the designed strategy cluster is denoted by b; by b_C if the cluster is of strategy C, and by b_D if it is of strategy D. Note that we frame the designed cluster with an agent playing the other strategy on both sides, such to make sure it cannot change size when the remaining strategies are drawn randomly. An example in a population of $N = 30$ displays the notation.

$$b_D = 3: \;\Rightarrow\; \underbrace{CDDDC}_{\text{fix}}\underbrace{CCDDDCCDDDCCDCCCDDCDDDCCD}_{\text{drawn randomly for every } r}$$

$$b_D = 8: \;\Rightarrow\; \underbrace{CDDDDDDDDC}_{\text{fix}}\underbrace{CCDCDCCCCDCCCDCCDDCC}_{\text{drawn randomly for every } r}$$

$$b_C = 10: \;\Rightarrow\; \underbrace{DCCCCCCCCCCD}_{\text{fix}}\underbrace{CCDCDDDDCDDDDDDCDD}_{\text{drawn randomly for every } r}$$

In the following we explain, how the random parts of the initial strategy states have to be drawn. Note that the probability of an agent to play C initially cannot be $\frac{1}{2}$ anymore, since we want to compare initial strategy states with the property $E[\sigma_{ISS}] = \frac{1}{2}$. We derive the following probabilities for agents in the random part of the initial strategy state:

$$b_C \text{ given} \;\Rightarrow\; \text{an agent plays } C \text{ with probability } \frac{\frac{N}{2} - b_C}{N - b_C - 2}$$

$$b_D \text{ given} \;\Rightarrow\; \text{an agent plays } C \text{ with probability } \frac{\frac{N}{2} - 2}{N - b_D - 2}.$$

The -2 in the above formulas are due to the fact that we enclose our fix part of the initial strategy state with two agents playing the other strategy. We also note that b_C and b_D are smaller than $\frac{N}{2}$.

For small b, we expect to receive the same results as in section 3.3, since the fix part in the initial strategy state most likely appears in a random strategy state anyway. It will therefore not influence the absorbing states. The higher b is though, the more different results we expect. Further we can also note the following. The higher b_C is, the larger the strategy clusters of D are that appear in the randomly drawn part of the initial strategy state, and vice versa. Thus we expect similar results for high b_C and b_D. We display the simulation results in Figure 3.6. We plot the share of agents playing C in the absorbing states that evolve from initial strategy states with different b. We state our findings from Figure 3.6 in the next result.

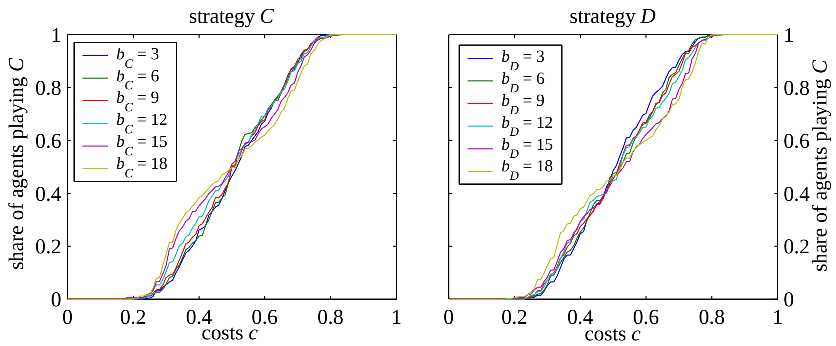


Figure 3.6: Strategy cluster in initial strategy states

Result 6 *If costs c are low, initial strategy states that exhibit a big cluster of strategy C (D), lead to absorbing states with a higher frequency of C than initial strategy states with small clusters only. For high costs c, the opposite is true.*

3.4.2 Varying the Type Distribution

We finally investigate how the ratio of H- and L-types affects the absorbing states. We define the ratio $h = \frac{n_H}{N}$ and draw $R = 500$ type distributions with the property $E[\frac{n_H}{N}] = h$. Simultaneously, initial strategy states are drawn randomly according to the standard settings. We then calculate the expected share of agents playing C in the absorbing states $\sigma^h(c)$. See Figure 3.7 for the results. We conclude the findings in the next result.

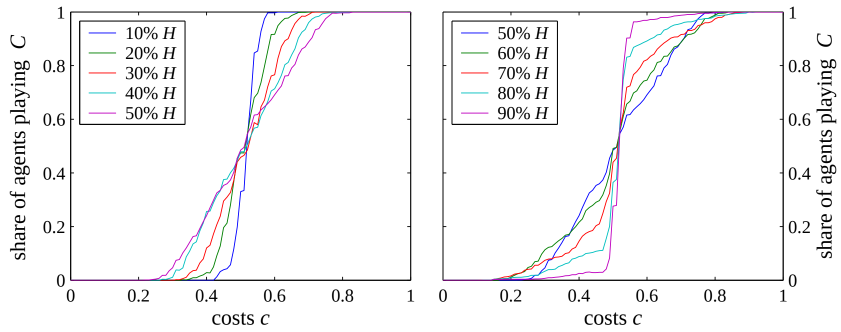


Figure 3.7: Different ratios of H- and L-types

Result 7 *The more the ratio $h = \frac{n_H}{N}$ diverges from $\frac{1}{2}$, the steeper $\sigma^h(c)$ is. For $c < \frac{1}{2}$ a ratio h close to $\frac{1}{2}$ yields the highest frequency of C, while for $c > \frac{1}{2}$ disproportionate numbers of types yield higher frequencies of C in the absorbing states.*

If we apply this result to policy implications, we can say that policy maker facing small costs c chooses a population that has equally many H- and L-types. If he faces high costs c though, a population with mainly H- or L-types will feature lower shares of C-playing agents.

3.5 Conclusions

In this chapter we take up the model of Chapter 2 in order to study the characteristics of its absorbing states under more general assumptions. We are concerned with the model variant that features local information and heterogenous agents. This type of model is ideally suited to analyze how illegal activities spread in a population. The reason is that agents making a decision wether to adopt an illegal activity or not, face two informational deficits: scarcity and non-verifiability (see 64, Chapter 2). These two informational deficits are comprehended by local information, heterogeneity, and imitation in our model. As a consequence, some agents under- or overestimate the value of the illegal activity (see 79, Chapter 2). We are interested in how this feature affects the absorbing states under general initial strategy states and type distributions.

Since the imitation dynamics of the model are analytically intractable when acting on the assumption of random initial strategy states and type distributions, we choose to simulate the model with a Monte Carlo method. This allows us to calculate the expected value of a given characteristic of an absorbing state. We are interested in three categories of characteristics.

The first category is interesting to interpret at the agent level. We find that high types exhibit a greater probability of acting illegally than low types, if the illegal activity bears small costs. If the illegal activity bears large costs, the low types are more likely to act illegally than the high types. We further find out, that agents surrounded only by types of their own type (interior players), display very different probabilities from agents that share information with agents of the other type (0-players). High types that are 0-players show very similar probabilities to adopt the illegal activity as low types that are 0-players. However, for interior players, the type does crucially matter for the probability to act illegally in the absorbing states.

The second category of characteristics concerns the sequence of the two strategies in the absorbing states. We find out that the absorbing states of our model exhibit strategy clusters, that are two to three times larger than those of a randomly drawn strategy state. The probability that an agent acts illegally if his neighbor does, is much larger in the model's absorbing states than in a random strategy state. We conclude that activities, both illegal and legal, exhibit a contagious character in our model.

The third category of characteristics is interesting if a policy maker is able to design either the initial strategy state or the distribution of types. An initial strategy state with a low share of agents acting illegally, leads to an absorbing state with a lower frequency of the illegal activity than an initial strategy state with a high share of agents acting illegally. If costs are low, initial strategy states with big strategy clusters lead to absorbing states with a lower frequency of the illegal activity than initial strategy states with little strategy clusters. The opposite is true for high costs.

The results of our model are evidence that social interactions offer explanations for the observed high variance of crime rates. We show that informational deficits can account for significant differences in the frequencies of illegal activities. Although our model includes features of social interactions and bounded rationality, it is still in line with the empirical fact, that high costs of an illegal activity, reduce its frequency.

3.A Appendix: An Example of the Evolutionary Game

We display an example of our evolutionary game here such that the reader gets an impression of how the imitation dynamics evolves from the initial strategy state to the absorbing state. To get a convenient table that can be included in the text, we choose a moderate population size of $N = 12$. In the top line of the table we display the type distribution which is drawn randomly. It is followed by the initial strategy state (labelled by $t = 1$) that also is drawn randomly. In the lines below we list the strategy state of every period. The agents change their strategies according to the imitation rule (2.3) and the payoffs they observe. The payoffs are listed below the respective strategy state. The strategy states are noted as the d_i (see the imitation rule in Chapter 2) of every agent: 0 stays for an agent playing C, while a 1 indicates a player that uses D. In the second table

TS	H	H	L	H	H	H	L	L	H	L	H	H
$t = 1$	0	0	0	0	1	1	0	0	1	0	0	1
	0.5	0.5	0.14	0.5	0.36	0.36	0.14	0.14	0.36	0.14	0.5	0.36
$t = 2$	0	0	0	1	0	1	1	1	1	1	1	0
	0.23	0.23	0	0.36	0.23	0.36	0.05	0.05	0.36	0.05	0.36	0.23
$t = 3$	0	0	1	1	1	0	1	1	1	1	0	1
	0.14	0.14	0	0.36	0.36	0.14	0	0	0.36	0	0.14	0.36
$t = 4$	1	0	1	1	1	1	0	1	1	1	1	1
	0.23	0.09	−0.23	0.23	0.23	0.23	0	−0.23	0.23	−0.23	0.23	0.23
$t = 5$	1	0	0	1	1	1	0	0	1	1	1	1
	0.23	0.27	0.09	0.23	0.23	0.23	0.09	0.09	0.23	−0.14	0.23	0.23
$t = 6$	0	1	1	1	1	1	1	1	0	1	1	1
	0.05	0.27	−0.18	0.27	0.27	0.27	−0.18	−0.18	0.05	−0.18	0.27	0.27
$t = 7$	1	1	1	1	1	1	1	0	0	1	1	1
	0.23	0.23	−0.23	0.23	0.23	0.23	−0.23	0	0.09	−0.23	0.23	0.23
$t = 8$	1	1	1	1	1	1	0	0	0	0	1	1
	0.23	0.23	−0.14	0.23	0.23	0.23	0.09	0.09	0.27	0.09	0.23	0.23
$t = 9$	1	1	1	1	1	1	1	0	0	1	1	1
	0.23	0.23	−0.23	0.23	0.23	0.23	−0.23	0	0.09	−0.23	0.23	0.23

we exclusively list the strategies so that one can easily see how the illegal technology evolves. We can check out all the features of the model that have been discussed in Chapter 2. First of all, let us look at a typical decision process. Agent 2 for instance cannot observe the technology in the first period and must play C in the following. In contrast, position 5 observes his left neighbor who plays C and gets the payoff 0.5, and his right neighbor who has a lower payoff with strategy D. From that, agent 5 concludes that it is not worth to keep playing D and plays C in period $t = 2$. Agent 4 and

TS	H	H	L	H	H	H	L	L	H	L	H	H
$t=1$	0	0	0	0	1	1	0	0	1	0	0	1
$t=2$	0	0	0	1	0	1	1	1	1	1	1	0
$t=3$	0	0	1	1	1	0	1	1	1	1	0	1
$t=4$	1	0	1	1	1	1	0	1	1	1	1	1
$t=5$	1	0	0	1	1	1	0	0	1	1	1	1
$t=6$	0	1	1	1	1	1	1	1	0	1	1	1
$t=7$	1	1	1	1	1	1	1	0	0	1	1	1
$t=8$	1	1	1	1	1	1	0	0	0	0	1	1
$t=9$	1	1	1	1	1	1	1	0	0	1	1	1

agent 11 both underestimate (see Chapter 2, Definition 3, on page 79) the strategy C and they do not choose the strategy that presently yields the highest payoff for them. Both of them underestimate C because they each have an L-type in their information set that does play C.

As we can see over the nine periods, there are many effects responsible for the absorbing state. Agents are under- or overestimate strategies, the environment changes too fast for the myopic agents to adopt, and the technology is not always available. Although we face complex dynamics where similar effects can have contradictory consequences, we show in this chapter that some positive statements about the absorbing states can be found.

Notation

Chapter 1

A	payoff matrix of basic corruption game
a_{hk}	elements of A
$\tilde{A}$	payoff matrix of basic corruption game
$\tilde{a}_{kl}$	elements of A
$A(x)$	payoff matrix of basic corruption game
α	parameter of distribution function Φ
β	parameter of distribution function Φ
c	individual cost of corruption in the basic corruption game
$c(x)$	strategy state dependent individual cost of corruption
ε	random variable with cumulative distribution function Φ
$f(\sigma, x)$	expected payoff function for strategy choice σ and strategy state x
F	right hand side of differential equation system defining the imitation dynamics
h	general index
i	strategy index
j	general index
k	general index
l	strategy state dependent legal population income
n	dimension of S
$\tilde{o}(x)$	expected payoff from corruption
$p(x)$	detection probability of a corrupt activity
$\varphi_i^j(x)$	probability that a player with strategy i switches to strategy j when reviewing his strategy
Φ	cumulative distribution function of random variable ε
Pr	probability

$r(x)$	tax revenue
$\rho_i(x)$	probability that a player with strategy i reviews his strategy choice at any given t; in the model we set $\rho_i(x) = 1$
s	surplus from private activity in basic corruption game
$\tilde{s}(x)$	strategy state dependent surplus from private activity before taxes
$s(x)$	strategy state dependent surplus from private activity after taxes
S	pure-strategy set of corruption game
σ	strategy choice of a player, a vector over S
Σ	simplex of dimension $n-1$, corresponds to strategy state space of corruption games
t	time, continuous
τ	tax rate
w	government wage in the basic corruption game
$\tilde{w}(x)$	strategy state dependent government wage before taxes
$w(x)$	strategy state dependent government wage after taxes
y^T	transpose of vector y
$x(t)$ or x	strategy state of the game (at time t), $x \in \Sigma_n$
$x_i(t)$ or x_i	frequency of strategy i (at time t)
$x_G(t)$ or x_G	frequency of public servants
x^0	strategy state in $t = 0$: initial condition
$\dot{y}(t)$ or $\dot{y}$	derivative of x with respect to t; if y a vector, derivatives are taken element-wise
y	a general variable

Chapter 2

A	payoff matrix of stage game for agents of the same type
$A_{H,L}$	payoff matrix of stage game for H-type against L-type
$A_{L,H}$	payoff matrix of stage game for L-type against H-type
$a(\sigma_{i,t}, \sigma_{j,t})$	corresponding element of A, given the strategy choices of agent i and agent j
c	individual cost of cheating
C	strategy: playing fair

$\vec{C}$	strategy state in which all agents play C
$\vec{C}\vec{C}$	strategy state: all H-types and all L-types play C
$\vec{C}\vec{D}$	strategy state: all H-types play C and all L-types D
$\vec{C}\vec{D}^*$	strategy state: identical to $\vec{C}\vec{D}$ except that L^E can play C
$\tilde{C}_{y_H}\tilde{C}_{y_L}$	strategy state: y_H H-types and y_L L-types play C, the agents playing C are all neighbors
$\tilde{C}_{y_H}\tilde{C}_{y_L}$	strategy state: identical to $\tilde{C}_{y_H}\tilde{C}_{y_L}$ except that two L^E can cycle between C and D
$d_{i,t}$	indicator variable: $d_{i,t} = 0$ if $\sigma_{i,t} = C$, $d_{i,t} = 1$ if $\sigma_{i,t} = D$
D	strategy: playing with illegal means, cheating
$\vec{D}$	strategy state in which all agents play D
$\vec{D}\vec{C}$	strategy state: all H-types play D and all L-types C
$\vec{D}\vec{C}^*$	strategy state: identical to $\vec{D}\vec{C}$ except that L^E can play D
$\vec{D}\vec{D}$	strategy state: all H-types and all L-types play D
$\Delta_{i,t}$	difference observed by agent i in period t between the average payoff of agents playing D and the average payoff of agents playing C
ε	probability of a mutation after strategy choice
$G_{i,t}(k)$	information set of agent i in period t
H	high type agent
H^E	H-type edge player
H^{EE}	H-type double-edge player
H^I	H-type interior player
H^N	H-type interior player who has an H^E as a neighbor
h	general index
i	specific agent
j	general index
k	size of information set
L	low type agent
L^E	L-type edge player
L^{EE}	L-type double-edge player
L^I	L-type interior player
L^N	L-type interior player who has an L^E as a neighbor
l	general index
M	set of stationary strategy states
μ	unique stationary probability distribution of Markov process

N	number of agents
n_H	number of H-types
n_L	number of L-types
n	equals $\frac{N}{2}$
$\lfloor\frac{N}{2}\rfloor$	largest integer smaller than $\frac{N}{2}$
ν	element of Σ_N
p_{ab}	probability that strategy state s_a is followed by s_b
P	transition probability matrix of Markov process
S	strategy state space
s	strategy state in period t
$\sigma_{i,t}$	strategy choice (either C or D) of agent i in period t
$\sigma_{-i,t}$	strategy choices of all but agent i in period t
Σ_N	2^N-dimensional simplex
t	period of time, discrete
$u_{i,t}$	agent i's payoff in period t
w	value of prize, normalized to 1 in the model
y	number of agents playing C (homogenous agents)
y_H	number of H-types playing C
y_L	number of L-types playing C
$z_{i,t}$	a random variable

Chapter 3

Notation of Chapter 2 is used in Chapter 3 too. Additionally, we have the following notation:

b_C	size of a C-strategy cluster
b_D	size of a D-strategy cluster
$\eta(c)$	average number of strategy clusters in absorbing state
$\eta^{pot}(c)$	maximal number of strategy clusters in absorbing state
$\eta^{ran}(c)$	random number of strategy clusters in a strategy state
g	position number of an agent concerning type state
γ	a characteristic of an absorbing state
h	expected share of H-types in a population
R	number of random realizations of stochastic process
$\sigma(c)$	average share of C-playing agents in absorbing state

$\sigma_H(c)$	average share of H-types playing C in absorbing state
$\sigma_L(c)$	average share of L-types playing C in absorbing state
$\sigma^g(c)$	average share of C-playing agents with position g in absorbing state
σ_{ISS}	share of agents playing C in initial strategy state
$\sigma^a(c)$	average share of agents playing C in absorbing state if in initial strategy state the share of C-playing agents is a

Bibliography

Ades, A. and R. Di Tella (1997). National champions and corruption: Some unpleasant interventionist arithmetic. *The Economic Journal 107*(443), 1023–42.

Alos-Ferrer, C., A. B. Ania, and K. R. Schenk-Hoppe (2000). An evolutionary model of bertrand oligopoly. *Games and Economic Behavior 33*(1), 1–19.

Amemiya, T. (1998). *Advanced Econometrics*. Harvard University Press.

Andvig, J. C. and K. O. Moene (1990). How corruption may corrupt. *Journal of Economic Behavior and Organization 13*(1), 63–76.

Bardhan, P. (1997, September). Corruption and development: A review of issues. *Journal of Economic Literature 35*(3), 1320–1346.

Barreto, R. A. (2000). Endogenous corruption in a neoclassical growth model. *European Economic Review 44*(1), 35–60.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Brenner, T. and U. Witt (2003). Melioration learning in games with constant and frequency-dependent payoffs. *Journal of Economic Behavior and Organization 50*(4), 429–448.

Brunetti, A., G. Kisunko, and B. Weder (1998). Credibility of rules and economic growth: Evidence from a worldwide survey of the private sector. *World Bank Economic Review 12*(3), 353–384.

Buscaglia, E. (2001). An analysis of judicial corruption and its causes: An objective governing-based approach. *International Review of Law and Economics 21*(2), 233–249.

Cartwright, E. (2002). Learning to play approximate nash equilibria in games with many players. Levine's Working Paper Archive, UCLA Department of Economics.

Chakrabarti, R. (2001). Corruption: A general equilibrium approach. Georgia Institute of Technology.

Chong, A. and C. Calderon (2000). Causality and feedback between institutional measures and economic growth. *Economics and Politics 12*(1), 69–81.

Dequech, D. (2001). Bounded rationality, institutions, and uncertainty. *Journal of Economic Issues 35*(4), 911–929.

Dollar, D. and A. Kraay (2003). Institutions, trade, and growth. *Journal of Monetary Economics 50*(1), 133–162.

Eggertsson, T. (1990). *Economic Behavior and Institutions.* Cambridge University Press.

Ehrlich, I. and F. T. Lui (1999). Bureaucratic corruption and endogenous economic growth. *Journal of Political Economy 107*(6), 270–293.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy 81*(3), 521–565.

Ellison, G. and D. Fudenberg (1995). Word-of-mouth communication and social learning. *Quarterly Journal of Economics 110*(1), 93–125.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica 61*(5), 1047–1071.

Eshel, I., L. Samuelson, and A. Shaked (1998). Altruists, egoists, and hooligans in a local interaction model. *American Economic Review 88*(1), 157–179.

Falk, A. and U. Fischbacher (2002). "crime" in the lab-detection social interaction. *European Economic Review 46*(4-5), 859–869.

Friedman, D. (1998). On economic applications of evolutionary game theory. *Journal of Evolutionary Economics 8*(1), 15–43.

Geweke, J. (1979). Measurement of linear dependence and feedback between time series. *Journal of the American Statistical Association 77*(378), 304–324.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996). Crime and social interactions. *The Quarterly Journal of Economics 111*(2), 507–548.

Goel, R. K. and M. A. Nelson (1998). Corruption and government size: A disaggregated analysis. *Public Choice 97*(1-2), 107–120.

Graf Lambsdorff, J. (1999). Corruption in empirical research - a review. Transparency International Working Paper.

Greene, W. H. (2000). *Econometric Analysis* (Forth ed.). Prentice-Hall, Inc.

Gupta, S., H. Davoodi, and R. Alonso-Terme (2002). Does corruption affect income inequality and poverty? *Economics of Governance 3*(1), 23–45.

Hewitt, D. and C. Van Rijckeghem (1995). Wage expenditures of central governments. International Monetary Fund Working Paper 9/11.

Hirshleifer, J. (1982). Evolutionary models in economics and law: Cooperative versus conflict strategies. *Research in Law and Economis 4*, 1–60.

Hofbauer, J. and K. Sigmund (1988). *The Theory of Evolution and Dynamical Systems.* Cambridge University Press.

Hofbauer, J. and K. Sigmund (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press.

Hofstaedter, R. (1973). *The American Political Tradition and the Men Who Made It.* Random House.

Husted, B. W. (1999). Wealth, culture, and corruption. *Journal of International Business Studies 30*(2), 339–359.

Hwang, J. (2002). A note on the relationship between corruption and government revenue. *Journal of Economic Development 27*(2), 161–177.

Jain, A. K. (2001). Corruption: A review. *Journal of Economic Surveys 15*(1), 71–121.

Joosten, R., T. Brenner, and U. Witt (2000). On games with frequency-dependent payoffs. Papers on Economics and Evolution #0003, Edited by the European Study Group for Evolutionary Economics.

Joosten, R., H. Peters, and F. Thuijsman (1994). Games with changing payoffs. In G. Silverberg and L. Soete (Eds.), *The Economics of Growth and Technical Change*, pp. 244–257. Aldershot, Edward Elgar.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT Press.

Kandori, M., G. J. Mailath, and R. Rob (1993). Learning, mutation, and long run equilibria in games. *Econometrica 61*(1), 29–56.

Kaufmann, D., A. Kraay, and P. Zoido-Lobatón (2000). Governance matters: From measurement to action. *Finance and Development 37*(2), 10–13.

Keefer, P. and S. Knack (1997). Why don't poor countries catch up? a cross-national test of institutional explanation. *Economic Inquiry 35*(3), 590–602.

Knack, S. and P. Keefer (1995). Institutions and economic performance: Cross-country tests using alternative institutional measures. *Economics and Politics 7*(3), 207–227.

Koenigsberger, K. (2001). *Analysis 1*. Springer.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Harvester Wheatsheaf.

Krueger, A. P. (1974). The political economcy of a rent-seeking society. *American Economic Review 64*(3), 291–303.

Lambsdorff, J. G. (2002). Making corrupt deals: Contracting in the shadow of the law. *Journal of Economic Behavior and Organization 48*(3), 221–241.

Li, H., L. C. Xu, and H. F. Zou (2000). Corruption, income distribution, and growth. *Economics and Politics 12*(2), 155–182.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives 14*(3), 115–136.

Mauro, P. (1995). Corruption and growth. *Quarterly Journal of Economics 110*(3), 681–712.

Mauro, P. (1997). The effects of corruption on growth, investment, and government expenditure: A cross-country analysis. In K. A. Elliott (Ed.), *Corruption and the Global Economy*, Chapter 4, pp. 83–107. Washington, D.C.: Institute for International Economics.

Mauro, P. (1998). Corruption and the composition of government expenditure. *Journal of Public Economics 69*(2), 263–279.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1993). Why is rent-seeking so costly to growth? *American Economic Review 83*(2), 409–414.

Niskanen, W. A. (1996). *Bureaucracy and Public Economics*. Edward Elgar Pub.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Ott, I. (2000). Bureaucratic choice and endogenous growth. *FinanzArchiv 57*(2), 225–242.

Paldam, M. (2002). The cross-country pattern of corruption: Economics, culture and the seesaw dynamics. *European Journal of Political Economy 18*(2), 215–240.

Perko, L. (2000). *Differential Equations and Dynamical Systems*. Springer.

Rose-Ackerman, S. (1978). *Corruption, A Study in Political Economy*. Academic Press.

Rose-Ackerman, S. (1999). *Corruption and Government*. Cambridge University Press.

Sah, R. K. (1991). Social osmosis and patterns of crime. *Journal of Political Economy 99*(6), 1272–1295.

Schuster, P. and K. Sigmund (1983). Replicator dynamics. *Journal of Theoretical Biology 100*, 533–538.

Selten, R. (2001). What is bounded rationality? In G. Gigerenzer and R. Selten (Eds.), *Bounded Rationality, The Adaptive Toolbox*, Chapter 2, pp. 13–37. The MIT Press.

Shleifer, A. and R. W. Vishny (1993). Corruption. *Quarterly Journal of Economics 108*(3), 599–617.

Simon, H. A. (1955). A beahvioral model of rational choice. *Quarterly Journal of Economics 69*, 99–118.

Simon, H. A. (1987). Bounded rationality. In J. Eatwell, M. Milgate, and P. Newman (Eds.), *The New Palgrave: A Dictionary of Economics*. Macmillan Publishers Limited.

Strogatz, S. H. (1994). *Nonlinear Dynamics and Chaos*. Westview Press, Perseus Books Group.

Tanzi, V. and H. R. Davoodi (2000). Corruption, growth, and public finances. International Monetary Fund Working Paper: WP/00/182 November 2000; 26.

Tanzi, V. (2002). Corruption around the world. *Applied Econometrics and International Development 2*(1), 110–115.

Taylor, P. D. and L. B. Jonker (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences 40*, 145–156.

Treisman, D. (2000). The causes of corruption: A cross-national study. *Journal of Public Economics 76*(3), 399–457.

Vega-Redondo, F. (1997). *Evolution, Games, and Economic Behaviour*. Oxford University Press.

Walter, W. (1991). *Analysis 2*. Springer-Verlag Berlin Heidelberg.

Weibull, J. W. (1995). *Evolutionary Game Theory*. The MIT Press.